NAVAL RESEARCH LOGISTICS QUARTERLY

JUNE 1973
VOL. 20, NO. 2

OFFICE OF NAVAL RESEARCH

NAVSO P-127

NAVAL RESEARCH LOGISTICS QUARTERLY

The Naval Research Logistics Quarterly is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Information for Contributors is indicated on inside back cover.

The Naval Research Logistics Quarterly is published by the Office of Naval Research in the months of March, June, September, and December and can be purchased from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Subscription Price: $10.00 a year in the U.S. and Canada, $12.50 elsewhere. Cost of individual issues may be obtained from the Superintendent of Documents.

Issuance of this periodical approved in accordance with Department of the Navy Publications and Printing Regulations, NAVEXOS P-35

OPTIMAL ORDERING POLICIES FOR A PRODUCT THAT PERISHES IN TWO PERIODS SUBJECT TO STOCHASTIC DEMAND*

Steven Nahmias

and

William P. Pierskalla

Northwestern University

Evanston, Illinois

ABSTRACT

This paper considers the problem of computing optimal ordering policies for a product that has a life of exactly two periods when demand is random. Initially costs are charged against runouts (stockouts) and outdating (perishing). By charging outdating costs according to the expected amount of outdating one period into the future, a feasible one period model is constructed. The central theorem deals with the n-stage dynamic problem and demonstrates the appropriate cost functions are convex in the decision variable and also provides bounds on certain derivatives. The model is then generalized to include ordering and holding costs. The paper is concluded with a discussion of the infinite horizon problem.

INTRODUCTION

Although some researchers feel that the single product inventory problem has been completely solved, there is one very important aspect to the problem which has received little attention in the literature; specifically, optimal ordering policies for a perishable product. Veinott [8] considered a somewhat restricted class of deterministic models in his dissertation. Bulinskaya [2] considered the case where the product perished at the end of the period in which it was ordered so that he was able to utilize the standard single product model. The most salient analysis done on this problem was that of Van Zyl [7]. The approach taken by Van Zyl was to charge only ordering and runout costs. Because of outdating, the average inventory entering each period would be less so that orders would be larger, thus in some sense accounting for perishing costs. Our approach will be to charge directly a cost for outdating so that reasonable one period results can be obtained and the cost of outdating can be made explicit.

Most of the effort involving a perishable product has been directed towards optimal issuing policies; however, from an applications point of view, the ordering problem is just as important. The farmer with the facility for stocking unsold produce is essentially faced with an ordering problem when deciding how much to plant. Hospitals that can order blood from a central blood bank face a similar problem. In addition, one can find applications of this problem to chemicals, radioactive elements, other aspects of the food processing industry, drugs, etc.

*This research was supported by the National Science Foundation under Grant GK–24792.

Our motivation in this paper is to develop a model for describing ordering policies for the product that perishes in two periods in the hope of introducing a structure that can be extended to solve the general problem. Wherever ordering occurs exactly twice during the product's useful life, this model will be applicable.

In the next section, the nature of the model is described and explicit statements of the assumptions are given. These assumptions are discussed in some detail and the probabilistic processes describing the perishability of the product are given.

In the third section, the existence of optimal ordering policies is established and many of its properties are demonstrated. Specialized results for the backlog case and the lost sales case are given as corollaries to the main theorems.

The fourth section is devoted to generalizing the results obtained to include ordering and holding costs as well. The introduction of a salvage value at the end of the horizon induces a type of stationarity on the model which is analogous to the single critical number obtained in Veinott [9].

In the last section, the results are extended to the discounted infinite horizon model.

THE NATURE OF THE MODEL—THE ONE PERIOD PROBLEM

Initially we will make the following assumptions. In the discussion which follows these assumptions, it will be pointed out which assumptions may be relaxed and which cannot be.

(1) All orders are placed at the start of the period and received instantaneously.

(2) All stock arrives new.

(3) Demands in each period are independent identically distributed nonnegative random variables with distribution function F and density f.

(4) Inventory is depleted according to a FIFO policy; that is, oldest first.

(5) Costs are charged linearly against

(a) unsatisfied demand (runouts) at r per unit

(b) deterioration (outdates) at θ per unit.

(6) If the product has not been depleted by the end of two periods then it deteriorates and must be discarded at a cost given in 5(b) above.

(7) Unsatisfied demand is backlogged.

Discussion of the Assumptions

(1) This assumption is standard for models which don't allow a delivery lag.

(2) This assumption is an important assumption for our model. As we will see, there is very little one can say when the stock that arrives is of mixed age.

(3) The fact that the demands are identically distributed is not an important assumption in the finite horizon problem. The assumption is made for notational convenience. It will be demonstrated that our results apply to nonstationary demand as well.

(4) This assumption is actually a consequence of the model and can easily be shown to be the optimal depletion policy. More general results of this type can be found in Pierskalla and Roach [5].

(5) We have assumed only two costs for the following reason: Consideration of a deterioration cost greatly complicates the model and for most perishable product problems deterioration and shortage costs are more important factors than holding and ordering costs. Again, though, this restriction is not crucial to the nature of the model and a more general cost structure will be considered.

(6) This assumption is the basic deterioration assumption.

(7) This assumption is not a crucial assumption. The model will also be applicable for the lost demand case, as will be indicated.

The state variable will be called x and will represent the amount of one period old product being brought into the period. The decision variable y represents the amount of new product being ordered. The first problem that must be considered is how to charge the costs. The obvious solution is charge costs against the expected cost incurred in the period so that

$$E\ [\text{Runouts}] = r\,E[D-(x+y)]^+ = r\int_{x+y}^{\infty}(t-(x+y))\,dF(t)$$

$$E\ [\text{Outdates}] = \theta\,E[x-D]^+ = \theta\int_0^x (x-t)\,dF(t),$$

where $f^+ = \max\,(f, 0)$ and F is the demand distribution. Letting

$$L(x, y) = r\int_{x+y}^{\infty}[(t-(x+y)]\ dF(t) + \theta\int_0^x (x-t)\ dF(t),$$

we see that this method has the property that y does not appear in the outdating cost, so that the charge for out-dating is actually independent of our order. Hence the value of y that minimizes $L(x, y)$ for any value of x is $y = +\infty$.

The direct approach to charging outdates leads to somewhat undesirable first period results so that we must develop a different method for charging the outdating cost.

Although there are undoubtedly many ways to charge the outdate cost, e.g., we could start the process with two periods to go, instead we will charge this cost to what the expected outdates will be after one period into the future. In this way, we will actually be charging the outdating cost against the expected outdating of the present order y. This approach has several advantages: first, it is reasonable for a real problem since the future demands are not known with certainty; second, it is easily generalizable to the case where a product perishes in m periods (rather than just two); third, it leads to a mathematically sensible one period problem as opposed to an approach which charges the outdate cost on the expected outdates $E[x-D]^+$ at the end of one period.

Let D_1, D_2 represent random variables denoting demands in two successive periods. Then the total amount of outdating of our present order will be

$$Z = \{y - [D_2 + (D_1 - x)^+]\}^+.$$

This can be seen most easily by considering the two possibilities $D_1 \leqslant x$ and $D_1 > x$ separately. If $D_1 \leqslant x$, then the demand in the first period is satisfied totally by one period old stock (x) and $x - D_1$ outdates, so that only the demand D_2 depletes from y, and when $D_2 \geqslant y$ there will be no outdating, and when $D_2 < y$, $y - D_2$ will outdate. In the other case, when $D_1 > x$, the total demand depleting from y will be $D_1 + D_2 - x$.

THEOREM 1: The random variable Z has distribution function H given by

$$H_{x,y}(t) = 1 - \int_0^{y-t} F(u+x)\,f(y-t-u)\,du \qquad t \geqslant 0$$

$$= 0 \qquad t < 0$$

and

$$E(Z) = \int_0^y F(u+x)\, F(y-u)\, du.$$

PROOF: We wish to compute

$$P\{(y - [D_2 + (D_1 - x)^+])^+ \leq t\} = P\{y - (D_2 + (D_1 - x)^+) \leq t\}$$

for $t \geq 0$, $t < y$.

$$P\{y - [D_2 + (D_1 - x)^+] \leq t\} = P\{y - (D_2 + (D_1 - x)^+) \leq t \cap D_1 \leq x\} + P\{y - (D_2 + (D_1 - x)^+) \leq t \cap D_1 > x\}$$

$$= P\{D_2 \geq y - t\} \cdot P\{D_1 \leq x\} + P\{D_1 + D_2 \geq y + x - t,\ D_1 > x\}$$

$$= [1 - F(y-t)]F(x) + \int_x^\infty [1 - F(y - t + x - u)]f(u)\, du$$

$$= F(x) - F(y-t)F(x) + (1 - F(x)) - \int_x^{y-t+x} F(y - t + x - u)f(u)\, du,$$

where

$$\int_x^{y-t+x} F(y - t + x - u)f(u)\, du = \int_0^{y-t} F(y - t - v)f(v + x)\, dv$$

$$= -F(y-t)F(x) + \int_0^{y-t} F(v + x)f(y - t - v)\, dv.$$

The last equality results from a change of orders of integration. Hence, substituting above we get

$$P\{y - [D_2 + (D_1 - x)^+]^+ \leq t\} = \begin{cases} 1 & \text{if} \quad t > y \\ 1 - \int_0^{y-t} F(u + x)f(y - t - u)\, du & 0 \leq t \leq y \\ 0 & t < 0. \end{cases}$$

Since Z is a nonnegative random variable it follows that

$$E(Z) = \int_0^\infty [1 - H_{x,y}(t)]\, dt = \int_0^y \int_0^{y-t} F(u + x)f(y - t - u)\, du\, dt$$

$$= \int_0^y \int_0^v F(u + x)f(v - u)\, du\, dv$$

$$= \int_0^y \int_v^y F(u+x)f(v-u)dv\,du = \int_0^y F(u+x)F(y-u)du. \qquad \text{Q.E.D.}$$

By utilizing this result, our one period expected cost function becomes

$$L(x, y) = r\int_{x+y}^{\infty} [t-(x+y)]f(t)dt + \theta\int_0^y F(t+x)F(y-t)dt.$$

Before examining the structure we've generated in detail, one can note a number of interesting properties immediately. Notice that $L(x, y)$ is not a function of $x+y$ as it is in the conventional models; the outdating cost differentiates between the inventory brought into the period and that ordered at the beginning of the period. Because of this property, the significant work on myopic policies by Veinott [8] and Bessler and Veinott [1] will not be applicable nor will single critical numbers or (s, S) policies be optimal.

THE DYNAMIC PROBLEM—FINITE HORIZON

The principle of optimality for this model takes the following form:

$$C_n(x) = \inf_{y\geq 0}\{L(x, y) + \alpha\int_0^{\infty} C_{n-1}(y-(t-x)^+)f(t)dt\},$$

where $C_n(x)$ has the usual interpretation of being the minimum expected cost given we have x on hand and there are n periods remaining in the horizon with $C_0(x)=0$ for all x. Note that we are discounting costs ($0<\alpha\leq 1$). The form of the transfer function $s(x, y, t)=y-(t-x)^+$ is a consequence of assumptions that we deplete according to a FIFO policy and demand is backlogged.

Define $B_n(x, y)=L(x, y)+\alpha\int_0^{\infty} C_{n-1}(y-(t-x)^+)f(t)dt$ and let $y_n(x)$ be such that

$$C_n(x) = B_n(x, y_n(x)) = \min_{y\geq 0}\{B_n(x, y)\}.$$

Then the purpose of the next theorem will be to demonstrate the existence of $y_n(x)$ and to enumerate some of its properties. The results obtained will be applicable to the finite horizon problem where the length of the horizon is arbitrary. Apologies are in order for the length of the theorem, but the induction argument requires all nine steps to be proven simultaneously.

THEOREM 2: Assuming that (a) the demand density $f(x)$ is continuous for all $x>0$, (b) $f(x)=0$ for all $x<0$, and (c) $f(x)>0$ for $x>0$. We then have

(1) $B_n(x, y)$ is convex in y for all fixed x and is strictly convex in y for a fixed x in a neighborhood of the global minimum.

(2) $\lim_{y\to 0}\frac{\partial B_n(x, y)}{\partial y}<0$ and $\lim_{y\to+\infty}\frac{\partial B_n(x, y)}{\partial y}>0$ for all x.

(3) There is a unique $y_n(x)$ given by the solution to $\frac{\partial B_n(x, y)}{\partial y}=0$, and $y_n(x)\in(0, \infty)$. In addition $\frac{dy_n(x)}{dx}=y_n'(x)$ exists and is continuous for all x.

(4) $C_n'(x)=-\alpha C_{n-1}'(y_n(x))F(x)-\theta F(y_n(x))F(x)$.

(5) $-1 \leq y_n'(x) < 0$. In addition $y_n'(x) > -1$ if $x > 0$.

(6) $C_n(x)$ is twice continuously differentiable over $(-\infty, 0)$ and $(0, \infty)$. $C_n''(x)$ is continuous at $x = 0$ whenever $f(x)$ is continuous at $x = 0$.

(7) $\theta f(x) + \alpha C_n''(x) \geq 0$ for all x.

(8) $-\theta F(x) \leq C_n'(x) \leq 0$ for all x and $C_n'(x) = 0$ for all $x \leq 0$. In addition $C_n'(x) < 0$ for $x \in (0, +\infty)$ and $n \geq 1$.

(9) $\lim_{x \to +\infty} y_n(x) = 0$ and $\lim_{x \to +\infty} C_n'(x) = 0$.

Before proving this theorem, it is useful to describe what some of these results mean. Parts (8) and (9) tell us that the optimal cost function, $C_n(x)$, is constant for $x \leq 0$ and then decreases to zero as $x \to +\infty$. Similarly part (5) indicates the optimal ordering function, $y_n(x)$, is decreasing in x, but at a rate greater than -1 and by part (9), this function also asymptotically approaches zero as $x \to +\infty$.

PROOF: The proof is by induction.

Assume the theorem is true for $1, 2, \ldots, n-1$.

The case of $n = 1$ follows in the same manner as when n is arbitrary with $C_0 \equiv 0$, so that the logic need not be repeated.

$$(1) \qquad B_n(x, y) = r\int_{x+y}^{\infty} [t - (x+y)] f(t)\,dt + \theta \int_0^y F(x+t)F(y-t)\,dt + \alpha \int_0^{\infty} C_{n-1}(y - (t-x)^+) f(t)\,dt$$

$$= r\int_{x+y}^{\infty} [t - (x+y)] f(t)\,dt + \theta \int_0^y F(x+t)F(y-t)\,dt + \alpha C_{n-1}(y)F(x) + \alpha \int_x^{\infty} C_{n-1}(x+y-t) f(t)\,dt$$

$$\frac{\partial B_n(x, y)}{\partial y} = -r[1 - F(x+y)] + \theta \int_0^y F(x+y-t) f(t)\,dt + \alpha C_{n-1}'(y)F(x) + \alpha \int_x^{x+y} C_{n-1}'(x+y-t) f(t)\,dt,$$

since by the inductive assumption on (6) and (8) C_{n-1}' exists and $C_{n-1}'(u) = 0$ for all $u \leq 0$.

$$(2) \qquad \frac{\partial^2 B_n(x, y)}{\partial y^2} = rf(x+y) + \theta F(x) f(y) + \theta \int_0^y f(x+y-t) f(t)\,dt + \alpha C_{n-1}''(y)F(x) + \alpha \int_x^{x+y} C_{n-1}''(x+y-t) f(t)\,dt$$

$$= rf(x+y) + F(x)[\theta f(y) + \alpha C_{n-1}''(y)] + \int_x^{x+y} [\theta f(x+y-t) + \alpha C_{n-1}''(x+y-t)] f(t)\,dt$$

$$\geq 0,$$

since by the inductive assumptions on (6) and (7) we have C_{n-1}'' exists and $\theta f(u) + C_{n-1}''(u) \geq 0$ for all u.

Note that in the case where f has a jump at 0, the differentiation under the integral sign in the third and fifth terms above would normally not be permitted. However, since the jumps occur at the endpoints of the range of integration and f and C_n'' are bounded the operation is valid. (See Van Zyl [7] for a lemma which establishes this.)

$$\lim_{y\to 0^+}\frac{\partial B_n(x,y)}{\partial y}=\lim_{y\to 0^+}\Big\{-r[1-F(x+y)]+\theta\int_0^y F(x+y-t)f(t)\,dt$$
$$+\alpha C'_{n-1}(y)F(x)+\alpha\int_x^{x+y}C'_{n-1}(x+y-t)f(t)\,dt\Big\}$$
$$=-r[1-F(x)]<0 \qquad \text{since } F(x)<1,$$

and by the inductive assumptions on (6) and (8) C'_{n-1} is continuous and $C'_{n-1}(0)=0$.

$$\lim_{y\to+\infty}\frac{\partial B_n(x,y)}{\partial y}=\theta-\alpha\lim_{y\to+\infty}\int_x^{x+y}C'_{n-1}(x+y-t)f(t)\,dt$$

by the inductive assumption on (9)

$$\geqslant\theta+\alpha\lim_{y\to\infty}\int_x^{x+y}-\theta F(x+y-t)f(t)\,dt,$$

by the inductive assumption on (8)

$$\geqslant\theta-\alpha\theta\lim_{y\to\infty}\int_x^{x+y}f(t)\,dt=\theta-\alpha\theta(1-F(x))\geqslant 0.$$

Thus $B_n(x,y)$ is convex and $\lim_{y\to 0^+}\frac{\partial B_n(x,y)}{\partial y}<0$ and $\lim_{y\to\infty}\frac{\partial B_n(x,y)}{\partial y}>0$ imply $y_n(x)\epsilon(0,+\infty)$, where $y_n(x)$ solves $\frac{\partial B_n(x,y)}{\partial y}=0$.

Now if $x<0$, then we must have $y_n(x)>|x|$ or else solving $\frac{\partial B_n(x,y)}{\partial y}=0$ we would obtain

$$-r[1-F(x+y_n(x))]+\theta\int_0^{y_n(x)}F(x+y_n(x)-t)f(t)\,dt+\alpha C_{n-1}(y_n(x))F(x)$$
$$+\alpha\int_x^{x+y_n(x)}C'_{n-1}(x+y_n(x)-t)f(t)\,dt=-r\neq 0,$$

which is impossible. Hence for all x, $f(x+y_n(x))>0$ and we see that in a neighborhood of $y_n(x)$

$$\frac{\partial^2 B_n(x,y)}{\partial y^2}>0$$

or $B_n(x,y)$ is strictly convex in y in a neighborhood of $y_n(x)$. Thus $y_n(x)$ is the unique global mini-

mum of $B_n(x, y)$. Define $T_n(x, y) = \dfrac{\partial B_n(x, y)}{\partial y}$ for all $(x, y) \in E^2$. Then

$$(3)\quad T_n(x, y) = -r[1 - F(x+y)] + \theta \int_0^y F(x+y-t)f(t)dt + \alpha C'_{n-1}(y)F(x)$$

$$+ \alpha \int_x^{x+y} C'_{n-1}(x+y-t)f(t)dt$$

is continuously differentiable on E^2 by the inductive assumption on (6). Now $(x, y_n(x))$ uniquely satisfies $T_n(x, y_n(x)) = 0$ and

$$\frac{\partial T_n(x, y_n(x))}{\partial y} = \frac{\partial^2 B_n(x, y_n(x))}{\partial y^2} > 0.$$

Hence by the implicit function theorem $y_n(x)$ is continuously differentiable in a neighborhood of x. But x was arbitrary in E^1 hence $y_n(x)$ is continuously differentiable for all x.

$$(4)\qquad C_n(x) = B_n(x, y_n(x))$$

$$\frac{dC_n(x)}{dx} = \frac{\partial B_n(x, y)}{\partial x}\bigg|_{y=y_n(x)} + \frac{\partial B_n(x, y)}{\partial y}\bigg|_{y=y_n(x)} \cdot y'_n(x)$$

$$= \frac{\partial B_n(x, y)}{\partial x}\bigg|_{y=y_n(x)} \text{ since } \frac{\partial B_n(x, y)}{\partial y}\bigg|_{y=y_n(x)} = 0.$$

Since $B_n(x, y_n(x)) = L(x, y_n(x)) + \alpha \int_0^\infty C_{n-1}[y_n(x) - (t-x)^+]f(t)dt$ it follows that

$$\frac{dC_n(x)}{dx} = \frac{\partial L(x, y)}{\partial x} + \alpha C_{n-1}(y)f(x) - \alpha C_{n-1}(y)f(x) + \alpha \int_x^\infty C'_{n-1}(x+y-t)f(t)dt\bigg|_{y=y_n(x)}$$

$$= -r[1 - F(x+y_n(x))] + \theta \int_0^{y_n(x)} f(x+y_n(x)-t)F(t)dt$$

$$+ \alpha \int_x^{x+y_n(x)} C'_{n-1}(x+y_n(x)-t)f(t)dt$$

by inductive assumption on (8).

But $y_n(x)$ satisfies

$$-r[1-F(x+y_n(x))] + \theta \int_0^{y_n(x)} F(x+y_n(x)-t)f(t)dt + \alpha C'_{n-1}(y_n(x))F(x)$$

$$+\alpha\int_x^{x+y_n(x)} C'_{n-1}(x+y_n(x)-t)f(t)\,dt=0.$$

Substituting, we obtain

$$\frac{dC_n(x)}{dx}=\theta\int_0^{y_n(x)} [f(x+y_n(x)-t)F(t)-F(x+y_n(x)-t)f(t)]\,dt-\alpha C'_{n-1}(y_n(x))F(x)$$

$$=-\theta F(y_n(x))F(x)-\alpha C'_{n-1}(y_n(x))F(x).$$

Since $y'_n(x)$ exists and since $T_n(x, y_n(x))=0$ for all x then

$$(5)\qquad \frac{\partial T_n(x,y_n(x))}{\partial x}=\frac{\partial^2 B_n(x,y_n(x))}{\partial x\,\partial y}$$

$$=rf(x+y_n(x))[1+y'_n(x)]+\theta y'_n(x)F(x)f(y_n(x))$$

$$-\theta\int_0^{y_n(x)} f(x+y_n(x)-t)[1+y'_n(x)]f(t)\,dt+\alpha C''_{n-1}(y_n(x))y'_n(x)F(x)$$

$$+\alpha\int_x^{x+y_n(x)} C''_{n-1}(x+y_n(x)-t)[1+y'_n(x)]f(t)\,dt,$$

by the inductive assumptions on (6) and (8) C''_{n-1} exists and $C'_{n-1}(0)=0$. Now since $y_n(x)>|x|$ if $x<0$ then $f(x+y_n(x))>0$ for all x; thus we can solve for $y'_n(x)$ to obtain

$$y'_n(x)=\frac{N(x,y_n(x))}{D(x,y_n(x))}\qquad \text{for all } x,$$

where

$$N(x,y_n(x))=-rf(x+y_n(x))+\int_x^{x+y_n(x)}[\theta f(x+y_n(x)-t)+\alpha C''_n(x+y_n(x)-t)]f(t)\,dt$$

<0 by inductive assumption on (7).

Also

$$D(x,y_n(x))=rf(x+y_n(x))+\theta F(x)f(y_n(x))+\int_x^{x+y_n(x)}[\theta f(x+y_n(x)-t)$$

$$+\alpha C''_{n-1}(x+y_n(x)-t)]f(t)\,dt+F(x)[\theta f(y_n(x))+\alpha C''_{n-1}(y_n(x))]$$

>0 by inductive assumption on (7).

Also since $-N(x,\ y_n(x))\leq D(x,\ y_n(x))$ it follows that $-1\leq y'_n(x)<0$. Notice that if $x>0$ the inequality is strict in both directions. We have that $y'_n(x)$ will exist and will be continuous by the inductive assumption on (6) (namely that C''_{n-1} exists and is continuous) so that

$$(6)\qquad C''_n(x)=\frac{d}{dx}[-\alpha C'_{n-1}(y_n(x))F(x)-\theta F(y_n(x))F(x)]$$

$$= -\alpha C''_{n-1}(y_n(x))y'_n(x)F(x) - \alpha C'_{n-1}(y_n(x))f(x)$$
$$-\theta f(y_n(x))y'_n(x)F(x) - \theta F(y_n(x))f(x)$$

is continuous.

(7) $$\theta f(x) + \alpha C''_n(x) = \theta f(x)[1 - \alpha F(y_n(x))] - \alpha^2 C'_{n-1}(y_n(x))f(x)$$
$$- y'_n(x)\alpha F(x)[\theta f(y_n(x)) + \alpha C''_{n-1}(y_n(x))]$$
$$\geqslant 0$$

since each term is nonnegative by the inductive assumptions on (7) and (8) and since $y'_n(x) < 0$.

(8) $$C'_n(x) = -\alpha C'_{n-1}(y_n(x))F(x) - \theta F(y_n(x))F(x) \geqslant -\theta F(y_n(x))F(x)$$

since $C'_{n-1}(x) \leqslant 0$, and

$$C'_n(x) \leqslant +\alpha\theta F(y_n(x))F(x) - \theta F(y_n(x))F(x) \leqslant 0$$

since $C'_{n-1}(x) \geqslant -\theta F(x)$.
Hence

$$-\theta F(x) \leqslant C'_n(x) \leqslant 0 \qquad \text{for all } x.$$

Assume to the contrary that $\lim_{x\to\infty} y_n(x) = w > 0$. (We know the limit exists since $y_n(x)$ is strictly decreasing and bounded below by zero.) Now we have for all x

(9) $$0 = \lim_{x\to\infty}\Big\{-r[1 - F(x + y_n(x))] + \theta\int_0^{y_n(x)} F(x + y_n(x) - t)f(t)\,dt$$
$$+ \alpha C'_{n-1}(y_n(x))F(x) + \alpha\int_x^{x+y_n(x)} C'_{n-1}(x + y_n(x) - t)f(t)\,dt\Big\}$$

$$= \alpha C'_{n-1}(w) + \lim_{x\to\infty}\Big\{+\theta\int_0^{y_n(x)} F(x + y_n(x) - t)f(t)\,dt + \alpha\int_x^{x+y_n(x)} C'_{n-1}(x + y_n(x) - t)f(t)\,dt\Big\}$$

$$\geqslant \alpha C'_{n-1}(w) + \theta\lim_{x\to\infty}\int_0^{y_n(x)} F(x)f(t)\,dt + \alpha\lim_{x\to\infty}\inf_{u\in E} C'_{n-1}(u)\int_x^{x+y_n(x)} f(t)\,dt$$

since $F(x) \leqslant F(x + y_n(x) - t)$ for all $t \in [0, y_n(x)]$.

By inductive assumptions (6), (8), and (9), since C'_{n-1} is continuous on E^1 and $\lim C'_{n-1}(x) = 0$ and $C'_{n-1}(x) = 0$ for all $x \leqslant 0$ then $\inf_{x\in E^1} C'_{n-1}(x) \equiv K > -\infty$. Thus

$$\liminf_{x\to\infty}{}_u\, C'_{n-1}(u)\int_x^{x+y_n(x)} f(t)\,dt = K\lim_{x\to\infty}\{F(x + y_n(x)) - F(x)\} = 0.$$

We see then from above that

$$0 \geqslant \alpha C'_{n-1}(w) + \theta F(w) \geqslant -\alpha\theta F(y_{n-1}(w))F(w) + \theta F(w) = \theta F(w)[1 - \alpha F(y_{n-1}(w))] > 0$$

via the inductive assumption on (8), and the fact that $w > 0$. Hence we achieve an impossibility, so it must be the case that

$$\lim_{x\to\infty} y_n(x) = 0.$$

Now

$$\lim_{x\to\infty} C_n'(x) = \lim_{x\to\infty} [-\alpha C_{n-1}'(y_n(x))\, F(x) - \theta F(y_n(x))\, F(x)]$$

$$= -\alpha C_{n-1}'(0) - \theta F(0) = 0. \qquad \text{Q.E.D.}$$

There are some very novel features of this model that are apparent from the proof. In particular, in the standard multi-period dynamic model for a nonperishable product the method of demonstrating convexity is quite different. One shows inductively that C_n is convex in its argument, which implies that $\int_0^\infty C_n\,[\underset{\sim}{s}\,(z,t)]\, f(t)dt$ is convex so that $c(z-x) + L(z) + \int_0^\infty C_n\,(\underset{\sim}{s}\,(z,t))\, f(t)dt$ is convex, where $L(z)$ is the expected holding and shortage costs when z is the inventory on hand after ordering. In our model through, C_n, in general, will not be convex so that it was necessary to use different methods to demonstrate the convexity of G_n.

The following corollary shows that when demand is backlogged, the optimal order quantity is a function of only a single critical number.

COROLLARY 1: If $x < 0$, then $y_n(x) = y_n(0) + |x|$.

PROOF: We know $y_n(x)$ satisfies

$$-r[1-F(x+y_n(x))] + \theta \int_0^{y_n(x)} F(x+y_n(x)-t) f(t)dt + \alpha \int_0^\infty C_{n-1}'(y_n(x) - (t-x)^+) f(t)dt = 0.$$

For $x < 0$ this becomes

$$-r(1-F(x+y_n(x))) + \theta \int_0^{y_n(x)+x} F(x+y_n(x)-t) f(t)dt + \alpha \int_0^\infty C_{n-1}'(y_n(x)+x-t) f(t)dt = 0.$$

Assume the solution is $y_n(x) = y_n(0) + |x| = y_n(0) - x$. Then by substituting in the above proof—

$$-r[1-F(y_n(0))] + \theta \int_0^{y_n(0)} F(y_n(0)-t) f(t)dt + \alpha \int_0^\infty C_{n-1}'(y_n(0)-t) f(t)dt.$$

But this is precisely the defining relationship for $y_n(0)$. Hence the substitution must necessarily have been valid. Q.E.D.

In the following corollaries we will indicate what generalizations apply.

COROLLARY 2: If the demand is nonstationary (where the indexing corresponds to the indexing of the functions C_n) and each of the demand densities satisfies the assumptions of Theorem 2, then all of the results of Theorem 2 and Corollary 1 remain valid with the following alterations:

(a) The recursive equations take the form

$$C_n(x) = \min_{y \geq 0} \left\{ L_n(x, y) + \alpha \int_0^\infty C_{n-1}(y - (t-x)^+) f_n(t)\, dt \right\},$$

where

$$L_n(x, y) = -r \int_{x+y}^\infty [t - (x+y)] f_n(t)\, dt + \theta \int_0^{y_n(x)} F_n(t+x) F_{n-1}(y-t)\, dt.$$

(b) Part (4) of the theorem becomes

$$C_n'(x) = -\alpha C_{n-1}'(y_n(x)) F_n(x) - \theta F_{n-1}(y_n(x)) F_n(x).$$

(c) Part (7) of the theorem becomes

$$\theta f_n(x) + \alpha C_n''(x) \geq 0 \qquad \text{for all } x.$$

(d) Part (8) of the theorem becomes

$$-\theta F_{n-1}(y_n(x)) F_n(x) \leq C_n'(x) \leq 0.$$

PROOF: (a) The functional equation form is well known. The form for the one period cost $L_n(x, y)$ is actually a consequence of Theorem 1. If we let the random variables be indexed backwards in accordance with the functional equations then from Theorem 1:

$$H_{x,y}(t) = P\{(y - [D_{n-1} + (D_n - x)^+]) \leq t\}$$

$$= \begin{cases} 1 - \int_0^{y-t} F_n(t+x) f_{n-1}(y-t)\, dt & t > 0 \\ 0 & t \leq 0, \end{cases}$$

so that E [outdates] $= \int_0^y F_n(t+x) F_{n-1}(y-t)\, dt.$

(b) Clearly (1)–(3) in Theorem 2 will not be affected by indexing the demands as long as each of the distributions, F_n, satisfies the assumptions of the theorem. To show (4) is true:

$$C_n(x) = L_n(x, y_n(x)) + \alpha \int_0^\infty C_{n-1}(y_n(x) - (t-x)^+) f_n(t)\, dt,$$

$$\frac{dC_n(x)}{dx} = \frac{dL_n(x, y_n(x))}{dx} + \alpha \int_x^{x+y_n(x)} C_{n-1}'(x + y_n(x) - t) f_n(t)\, dt,$$

where $y_n(x)$ satisfies

$$-r[1-F_n(x+y_n(x))]+\theta\int_0^{y_n(x)}F_n(t+x)f_{n-1}(y_n(x)-t)dt+\alpha C'_{n-1}(y_n(x))F_n(x)$$
$$+\alpha\int_x^{x+y_n(x)}C'_{n-1}(x+y_n(x)-t)f_n(t)dt=0.$$

Since

$$\frac{dL_n(x,\,y_n(x))}{dx}=-r[1-F_n(x+y_n(x))]+\theta\int_0^{y_n(x)}f_n(t+x)F_{n-1}(y_n(x)-t)dt,$$

it follows that

$$\frac{dC_n(x)}{dx}=\theta\int_0^{y_n(x)}[f_n(t+x)F_{n-1}(y_n(x)-t)-F_n(t+x)f_{n-1}(y_n(x)-t)]dt-\alpha C'_n(y_n(x))F_n(x)$$

$$=-\alpha C'_n(y_n(x))F_n(x)-\theta F_{n-1}(y_n(x))F_n(x)$$

so that (4) holds; (5) and (6) will remain valid in spite of the nonstationary demands. For (7) we will get

$$\alpha C''_n(x)+\theta f_n(x)=\theta f_n(x)[1-\alpha F_{n-1}(y_n(x))]-\alpha^2C'_{n-1}[y_n(x)]f_n(x)$$
$$-\alpha y'_n(x)F_n(x)[\theta f_{n-1}(y_n(x))+\alpha C''_{n-1}(y_n(x))]\geqslant 0.$$

Again the first term is positive; $C'_{n-1}(u)\leqslant 0$ for all u so that the second term is positive, and the inductive assumption on (7) yields the positivity of the last term.

(d) The generalization follows directly from (b) above.

The remaining sections of Theorem 2 remain valid under nonstationary demand.

Q.E.D.

Because of our method of projecting the outdate costs into the future, we must assume that the demand distribution, F_0, corresponding to the period after the end of the horizon is known (recall that periods are numbered backwards).

COROLLARY 3: If we do not allow the demand to be backlogged (lost sales case), where the transfer function takes the form

$$s(x,y,t)=(y-(t-x)^+)^+,$$

then all of the results of Theorem 2 remain valid (where $x\geqslant 0$ always).

PROOF: For $n=1$ the two problems are identical, so all results will follow. Assume all the results hold for $n=1$. Then

$$B_n(x,y)=r\int_{x+y}^{\infty}[t-(x+y)]f(t)dt+\theta\int_0^{y}F(x+t)\,F(y-t)\,dt+\alpha\int_0^{\infty}C_{n-1}[(y-(t-x)^+)^+]f(t)dt$$

$$=r\int_{x+y}^{\infty}[t-(x+y)]f(t)dt+\theta\int_0^{y}F(x+y-t)\,F(t)dt+\alpha C_{n-1}(y)\,F(x)$$
$$+\alpha\int_x^{x+y}C_{n-1}(y+x-t)f(t)dt+\alpha C_{n-1}(0)[1-F(x+y)],$$

$$\frac{\partial B_n(x,y)}{\partial y} = -r[1-F(x+y)] + \theta\int_0^y F(x+y-t)f(t)dt + \alpha C'_{n-1}(y)F(x)$$
$$+ \alpha\int_x^{x+y} C'_{n-1}(y+x-t)f(t)dt,$$

which is precisely the same form as in the backlogging case, so that (1), (2), and (3) will follow in the same way. To show (4) holds as well:

$$C_n(x) = \min_{y\geq 0}\{L(x,y) + \alpha\int_0^\infty C_{n-1}[(y-(t-x)^+)^+]f(t)dt\}$$
$$= L(x, y_n(x)) + \alpha C_{n-1}(y_n(x))F(x) + \alpha\int_x^{x+y_n(x)} C_{n-1}(y_n(x)+x-t)f(t)dt$$
$$+ \alpha C_{n-1}(0)[1-F(x+y_n(x))],$$

$$C'_n(x) = -r[1-F(x+y_n(x))] + \theta\int_0^{y_n(x)} f(x+y_n(x)-t)F(t)dt + \alpha\int_x^{x+y_n(x)} C'_{n-1}(y_n(x)+x-t)f(t)dt,$$

since all terms involving $y'_n(x)$ drop out. But $y_n(x)$ satisfies:

$$-r[1-F(x+y_n(x))] + \theta\int_0^{y_n(x)} F(x+y_n(x)-t)f(t)dt + \alpha C'_{n-1}(y_n(x))F(x)$$
$$+ \alpha\int_x^{x+y_n(x)} C'_{n-1}(x+y_n(x)-t)f(t)dt = 0.$$

Hence

$$C'_n(x) = \theta\int_0^{y_n(x)} [f(x+y_n(x)-t)F(t) - F(x+y_n(x)-t)f(t)]dt - \alpha C'_{n-1}(y_n(x))F(x)$$
$$= -\theta F(y_n(x))F(x) - \alpha C'_{n-1}(y_n(x))F(x)$$

as in Theorem 2.

Since the defining relationship for $y_n(x)$ is precisely the same in both cases, $y'_n(x)$ is also equivalent and (5) will hold. The results (6), (7), and (8) follow from (4), and (9) and (10) hold independent of backlogging assumptions. Note that all the results for this corollary only hold for $x \geq 0$. Q.E.D.

ADDITION OF HOLDING AND ORDERING COSTS

An important consideration in our model is whether or not the results can be extended to a more general cost structure including holding and ordering costs. The answer is that it can, but the form of the policy changes; in particular, it will not always be advantageous to place an order.

Assume that there is a charge of $c > 0$ per unit for each unit ordered and a charge of $h > 0$ per unit (charged at the end of the period before outdating, but after demands occur) for each unit carried. Hence the one period expected cost function now becomes

$$L(x, y) = cy + h \int_0^{x+y} (x+y-t)f(t)dt + r \int_{x+y}^{\infty} (t-(x+y))f(t)dt + \theta \int_0^{y} F(v+x)F(y-v)dv.$$

We will also make the assumption that there is a salvage value at the end of the horizon which is equivalent to the ordering cost so that

$$C_1(x) = \inf_{y \geq 0} \left\{ L(x, y) - \alpha c \int_0^{\infty} \{y-(t-x)^+\} f(t)dt \right\}.$$

It turns out that this assumption allows us to obtain a certain type of stationarity in our results. Notice that if we define

$$C_0(x) = -cx \qquad \text{for all } x,$$

then the functional equations are still consistently defined.

We will use the notation C_n and B_n as in Theorem 2. Also the same assumptions are made on the demand distribution. We also assume that $r > (1-\alpha)c$ (otherwise we would never order to a positive level).

We then have the following analogous result to Theorem 2:

THEOREM 3: Assume F has the same properties as in Theorem 2 and that $r > c(1-\alpha)$. Define the number $\bar{x}$ as the unique positive solution to

$$(1-\alpha)c + hF(\bar{x}) - r[1-F(\bar{x})] = 0, \text{ i.e., } \left[\bar{x} = F^{-1} \left[\frac{r-(1-\alpha)c}{r+h} \right] \right].$$

Then we have

(1) $B_n(x, y)$ is convex in y for all x, and is strictly convex in the neighborhood of the global minimum for $x \leq \bar{x}$.

(2) $\lim_{y \to 0} \frac{\partial B_n(x, y)}{\partial y} < 0$ if and only if $x < \bar{x}$, $\lim_{y \to +\infty} \frac{\partial B_n(x, y)}{\partial y} > 0$ for all x.

(3) If $x < \bar{x}$ then there exists a unique nonnegative solution $y_n(x)$ which solves

$$\left. \frac{\partial B_n(x, y)}{\partial y} \right|_{y=y_n(x)} = 0.$$

In addition $y_n'(x)$ exists and is continuous for $x < \bar{x}$.

If $x \geq \bar{x}$, then the optimal policy is not to order, i.e., $y_n(x) = 0$.

(4) If $x < \bar{x}$ then

$$C_n'(x) = -c - \alpha C_{n-1}'[y_n(x)]F(x) - \theta F(y_n(x))F(x).$$

(5), (6), and (7) of Theorem 2 hold for $x < \bar{x}$.

(8) If $0 < x < \bar{x}$, then $-c-\theta F(x) \leq C_n'(x) \leq 0$ and $C_n'(x) = -c$ if $x \leq 0$.

(9) If $x > \bar{x}$ then $C_n''(x) > 0$. $C_n'(x) > -c$. In addition C_n and C_n' are continuous at $\bar{x}$.

PROOF: Let $n=1$.

(1), (2), and (3): By assumption we obtain

$$B_1(x, y) = L(x, y) - \alpha c \int_0^\infty (y-(t-x)^+) f(t)\,dt$$

$$= cy + h\int_0^{x+y} (x+y-t) f(t)\,dt + r\int_{x+y}^\infty (t-(x+y)) f(t)\,dt + \theta \int_0^y F(v+x)F(y-v)\,dv$$

$$- \alpha c\left[yF(x) + \int_x^\infty (y+x-t) f(t)\,dt \right],$$

$$\frac{\partial B_1(x, y)}{\partial y} = c - \alpha c + hF(x+y) - r[1-F(x+y)] + \theta \int_0^y F(v+x) f(y-v)\,dv,$$

and

$$\frac{\partial^2 B_1(x, y)}{\partial y^2} = (h+r) f(x+y) + \theta F(x) f(y) + \theta \int_0^y f(v+x) f(y-v)\,dv \geq 0.$$

Hence

$$\left.\frac{\partial B_1(x, y)}{\partial y}\right|_{y=0} = c(1-\alpha) + hF(x) - r[1-F(x)] = c(1-\alpha) - r + (h+r) F(x).$$

Now since $\bar{x}$ is defined where $c(1-\alpha) + hF(\bar{x}) - r[1-F(\bar{x})] = 0$ and $\left.\frac{\partial B_1(x, y)}{\partial y}\right|_{y=0}$ is a monotone nondecreasing function of x (and is strictly increasing in a neighborhood of $\bar{x}$), it follows that

$$\left.\frac{\partial B_1(x, y)}{\partial y}\right|_{y=0} < 0 \qquad \text{if } x < \bar{x},$$

$$\left.\frac{\partial B_1(x, y)}{\partial y}\right|_{y=0} > 0 \qquad \text{if } x > \bar{x},$$

and

$$\lim_{y\to\infty} \frac{\partial B_1(x, y)}{\partial y} \geq (1-\alpha) c + h + \theta > 0 \text{ for all } x.$$

Hence it follows that if $x < \bar{x}$, the optimal policy is to order to $y_1(x)$ where $y_1(x)$ solves

$$\left.\frac{\partial B_1(x, y)}{\partial y}\right|_{y=y_1(x)} = 0,$$

and if $x > \bar{x}$ we don't order. Notice if $x = \bar{x}$ then $y_1(\bar{x}) = 0$, so the policy is continuous at $\bar{x}$.

The differentiability of $y_1(x)$ follows in precisely the same manner as in Theorem 2 when $x < \bar{x}$.

(4) If $x < \bar{x}$ then

$$C_1(x) = B_1(x, y_1(x))$$

$$C_1'(x) = \frac{\partial B_1(x, y)}{\partial x}\bigg|_{y=y_1(x)} + \frac{\partial B_1(x, y)}{\partial y}\bigg|_{y=y_1(x)} \cdot y_1'(x) = \frac{\partial B_1(x, y)}{\partial x}\bigg|_{y=y_1(x)}$$

$$= hF(x + y_1(x)) - r[1 - F(x + y_1(x))] + \theta \int_0^{y_1(x)} f(v + x) F(y_1(x) - v)\, dv - \alpha c(1 - F(x))$$

$$= \{c(1 - \alpha) + h[F(x + y_1(x))] - r[1 - F(x + y_1(x))]$$

$$+ \theta \int_0^{y_1(x)} F(v + x) f(y_1(x) - v)\, dv\Big\} - c - \theta F(x) F(y_1(x)) + \alpha c F(x)$$

$$= -c - \theta F(x) F(y_1(x)) + \alpha c F(x).$$

The term in brackets is

$$\text{exactly } \frac{\partial B_1(x, y)}{\partial y}\bigg|_{y=y_1(x)} = 0.$$

Also notice that $C_0'(x) = -c$, so this conforms with our notation.

(5) Solving for $y_1'(x)$, we obtain (as long as $x < \bar{x}$) precisely the same expression as in Theorem 2 with r replaced by $r + h$. All terms involving c drop out.

$$\lim_{x \uparrow \bar{x}} y_1'(x) = \frac{-(r + h) f(\bar{x})}{(r + h) f(\bar{x}) + \theta F(\bar{x}) f(0)} < 0$$

so that there is a discontinuity of y_1' at $\bar{x}$, since $y_1'(x) = 0$, for $x > \bar{x}$. Notice if $f(0) = 0$, then $\lim_{x \uparrow \bar{x}} y_1'(x) = -1$

(6) and (7) follow since C_1'' is identical. Notice that C_1'' will not be continuous at $\bar{x}$ by the fact that $y_1'(x)$ is not continuous at $\bar{x}$.

(8) follows directly from (4).

(9) If $x > \bar{x}$, $C_1(x) = B_1(x, 0)$.

$$C_1'(x) = hF(x) - r[1 - F(x)] - \alpha c[1 - F(x)]$$

$$= (1 - \alpha)c + hF(x) - r[1 - F(x)] - c + \alpha c F(x) > -c$$

$$C_1''(x) = (h + r + \alpha c) f(x) > 0 \text{ since } x > \bar{x} > 0.$$

Notice $\lim_{x \downarrow \bar{x}} C_1(x) = B_1(\bar{x}, 0) = \lim_{x \uparrow \bar{x}} C_1(x)$ and $\lim_{x \downarrow \bar{x}} C_1'(x) = \lim_{x \uparrow \bar{x}} C_1'(x) = -c + \alpha c F(\bar{x})$.

Assume that the theorem holds for $1, 2, \ldots, n-1$.

(1), (2), and (3):

$$B_n(x, y) = L(x, y) + \alpha \int_0^\infty C_{n-1}(y-(t-x)^+) f(t)\,dt.$$

Let $y < \bar{x}$.

$$\frac{\partial B_n(x, y)}{\partial y} = c + hF(x+y) - r[1-F(x+y)] + \theta \int_0^y F(v+x) f(y-v)\,dv$$
$$+ \alpha C'_{n-1}(y) F(x) + \alpha \int_x^{x+y} C'_{n-1}(y+x-t) f(t)\,dt - \alpha c[1-F(x-y)].$$

$$\frac{\partial^2 B_n(x, y)}{\partial y^2} = (h+r) f(x+y) + \theta F(x) f(y) + \theta \int_0^y f(v+x) f(y-v)\,dv + \alpha C''_n(y) F(x)$$
$$+ \int_x^{x+y} C''_{n-1}(y+x-t) f(t)\,dt + \alpha c f(x+y) - \alpha c f(x+y)$$

$$\geqslant (h+r) f(x+y) + \theta(1-\alpha) f(y) + \theta(1-\alpha) \int_0^y f(v+x) f(y-v)\,dv$$

$$\geqslant 0 \text{ by the inductive assumption on (7).}$$

If $y > \bar{x}$, then

$$B_n(x, y) = L(x, y) + \alpha C_{n-1}(y) F(x) + \alpha \int_x^{y+x-\bar{x}} C_{n-1}(y+x-t) f(t)\,dt + \alpha \int_{y+x-\bar{x}}^\infty C_{n-1}(y+x-t) f(t)\,dt.$$

$$\frac{\partial B_n(x, y)}{\partial y} = c + hF(x+y) - r[1-F(x+y)] + \theta \int_0^y F(v+x) f(y-v)$$
$$+ \alpha C'_{n-1}(y) F(x) + \alpha \int_x^{y+x-\bar{x}} C'_{n-1}(y+x-t) f(t)\,dt + \alpha C_{n-1}(\bar{x}) f(y+x-\bar{x}) - \alpha C_{n-1}(\bar{x}) f(y+x-\bar{x})$$
$$+ \alpha \int_{y+x-\bar{x}}^{y+x} C'_{n-1}(y+x-t) f(t)\,dt - \alpha c[1-F(y+x)].$$

Notice that the upper limit on the last integral became $y+x$ by the inductive assumption on (8).

$$\frac{\partial^2 B_n(x, y)}{\partial y^2} = (h+r) f(x+y) + \theta F(x) f(y) + \theta \int_0^y f(v+x) f(y-v)\,dv$$
$$+ \alpha C''_{n-1}(y) F(x) + \alpha \int_x^{y+x-\bar{x}} C''_{n-1}(y+x-t) f(t)\,dt + \alpha C'_{n-1}(\bar{x}) f(y+x-\bar{x})$$
$$- \alpha C'_{n-1}(\bar{x}) f(y+x-\bar{x}) + \alpha \int_{y+x-\bar{x}}^{y+x} C''_{n-1}(y+x-t) f(t)\,dt - \alpha c f(y+x) + \alpha c f(y+x).$$

Although C''_{n-1} may have a jump at $\bar{x}$ and at 0 the differentiation under the integral sign is justified by Van Zyl's lemma.

Now since $y > \bar{x}$, $C''_{n-1}(y) > 0$ by the inductive assumption on (9),

$$\int_x^{y+x-\bar{x}} C''_{n-1}(y+x-t)\, f(t)\, dt = \int_{\bar{x}}^{y} C''_{n-1}(u)\, f(y+x-u)\, du > 0$$

and in the same way,

$$\int_{y+x-\bar{x}}^{y+x} C''_{n-1}(y+x-t)\, f(t)\, dt = \int_0^{\bar{x}} C''_{n-1}(u)\, f(y+x-u)\, du$$

$$\geqslant -\theta \int_0^{\bar{x}} f(u)\, f(y+x-u)\, du$$

$$\geqslant -\theta \int_0^{y} f(u)\, f(y+x-u)\, du$$

by the inductive assumption on (7).

Hence

$$\frac{\partial^2 B_n(x,y)}{\partial y^2} > (h+r) f(x+y) + \theta F(x) f(y) + \theta(1-\alpha) \int_0^y f(v+x) f(y-v)\, dv \geqslant 0$$

and convexity is established over $(\bar{x}, \infty)$.

Since $B_n(x,y)$ and $\dfrac{\partial B_n(x,y)}{\partial y}$ are continuous at $\bar{x}$ the convexity follows.

$$\lim_{y\to 0} \frac{\partial B_n(x,y)}{\partial y} = c + hF(x) - r[1-F(x)] + \alpha \lim_{y\to 0} C'_{n-1}(y) F(x) + \alpha \int_x^{\infty} \lim_{y\to 0} C'_{n-1}(y+x-t) f(t)\, dt$$

$$= c(1-\alpha) + hF(x) - r[1-F(x)]$$

by the inductive assumption on (8) so that

$$\left.\frac{\partial B_n(\bar{x},y)}{\partial y}\right|_{y=0} = 0$$

and it follows that if $x < \bar{x}$, $\left.\dfrac{\partial B_n(x,y)}{\partial y}\right|_{y=0} < 0$ and the optimal policy is to order $y_n(x)$ where $y_n(x)$ is the solution to

$$\left.\frac{\partial B_n(x,y)}{\partial y}\right|_{y=y_n(x)} = 0$$

and if $x > \bar{x}$, $\left.\frac{\partial B_n(x,y)}{\partial y}\right|_{y=0} > 0$, so it is advantageous not to order. Notice $y_n(\bar{x}) = 0$.

(4) If $x < \bar{x}$, then

$$C_n'(x) = hF(x+y_n(x)) - r[1-F(x+y_n(x)] + \theta \int_0^{y_n(x)} f(v+x)F(y_n(x)-v)dv + \alpha f(x)C_{n-1}(y_n(x))$$

$$- \alpha f(x)C_{n-1}(y_n(x)) + \alpha \int_x^{\infty} C_{n-1}'(y_n(x)+x-t)f(t)dt$$

$$= \Big\{c + hF(x+y_n(x)) - r[1-F(x+y_n(x)] + \theta \int_0^{y_n(x)} F(v+x)f(y_n(x)-v)dv$$

$$+ \alpha F(x)C_{n-1}'(y_n(x)) + \alpha \int_x^{\infty} C_{n-1}'(y_n(x)+x-t)f(t)dt\Big\} - c - \alpha F(x)C_{n-1}(y_n(x))$$

$$- \theta F(y_n(x))F(x)$$

$$= -c - \alpha F(x)C_{n-1}(y_n(x)) - \theta F(y_n(x))F(x)$$

by the definition of $y_n(x)$.

(5), (6), and (7) follow in precisely the same way as in Theorem 2 since the expressions for $y_n'(x)$ and $C_n''(x)$ are the same for $x < \bar{x}$.

Notice that $\lim_{x \uparrow \bar{x}} y_n'(x) < 0$ and $= -1$ if $f(0) = 0$, as in the $n = 1$ case.

(8) From (4) and the inductive assumption on (8), we obtain for $x < \bar{x}$

$$C_n'(x) \leq -c - \alpha F(x)[-c - \theta F(y_n(x)] - \theta F(y_n(x)F(x)$$

$$\leq -c(1-\alpha) - \theta(1-\alpha)F(y_n(x))F(x) \leq 0,$$

$$C_n'(x) \geq -c - \theta F(y_n(x))F(x) \geq -c - \theta F(x).$$

From (4) $C_n'(x) = -c$ if $x \leq 0$.

(9) If $x > \bar{x}$ then

$$C_n(x) = L(x,0) + \alpha \int_0^{\infty} C_{n-1}(-(t-x)^+)f(t)dt$$

$$= L(x,0) + \alpha C_{n-1}(0)F(x) + \alpha \int_x^{\infty} C_{n-1}(x-t)f(t)dt.$$

Hence

$$C_n'(x) = hF(x) - r[1-F(x)] - \alpha c[1-F(x)]$$

$$> (1-\alpha)c + hF(\bar{x}) - r[1-F(\bar{x})] - c + \alpha cF(x) > -c$$

$$C_n''(x) = (h+r+\alpha c)f(x) > 0.$$

The continuity of $C_n(x)$ and $C_n'(x)$ at $\bar{x}$ follows easily since $y_n(\bar{x}) = 0$ by (1), (2), and (3), and $C_{n-1}'(0) = -c$ by (8). Thus $C_n'(\bar{x}) = -c + \alpha cF(\bar{x})$ independent of n. Q.E.D.

Notice that the introduction of a salvage value at the end of the horizon insures that the numbers $\bar{x}$ don't depend on the period n. If we exclude this assumption then $\bar{x}_2 > \bar{x}_1$ and $\bar{x}_n = \bar{x}_{n-1}$, $2 < n < N$, provided we number the periods backwards.

We remark that the results of Corollaries 1, 2, and 3 remain valid with the addition of ordering and holding costs.

The model considered by Van Zyl is precisely this one with $h = \theta = 0$. Note that in the case where $\theta = 0$ the results of the theorem remain valid with the added property that C_n is convex everywhere.

We have the following result:

COROLLARY 4: For $0 < x < \bar{x}$ and all n,

$$0 < y_n(0) < y_n(x) + x < \bar{x}.$$

PROOF: Since $y_n'(x) > -1$ and $y_n(x)$ is continuous,

$$-1 < \frac{y_n(x) - y_n(0)}{x} \text{ from the mean value theorem}$$

$$=> y_n(x) + x > y_n(0).$$

Also since $y_n'(x) > -1$ for $0 < x < \bar{x}$, $\lim_{x \uparrow \bar{x}} y_n'(x) \geq -1$ and $y_n(x)$ is continuous,

$$-1 < \frac{y_n(\bar{x}) - y_n(x)}{\bar{x} - x}$$

$$=> x + y_n(x) < y_n(\bar{x}) + \bar{x} = \bar{x}.$$ Q.E.D.

This is a particularly significant result. It says that if the initial inventory is less than $\bar{x}$ then we would order in every period. If $x > \bar{x}$, then we would not order until the inventory on hand fell below $\bar{x}$ and then we would order in each subsequent period.

In the case where $y_n(0)$ and $\bar{x}$ are relatively close, one could approximate an optimal ordering policy by assuming the inventory on hand after ordering is some number interpolated between $y_n(0)$ and $\bar{x}$.

THE DYNAMIC PROBLEM—INFINITE HORIZON

The objective in analyzing the infinite horizon problem is to demonstrate the existence of a stationary policy that minimizes the long run expected discounted cost and to determine methods for computing that policy. The formalism we will employ is that of Denardo [3].

The state space of our problem is $R = (-\infty, \infty)$.

Define:

$\mathcal{W} = \{w : w$ is a real valued Borel measurable function on $R\}$. The metric on $\mathcal{W}$ is defined in the usual way by

$$\rho(u, v) = \sup_{x \in R} | u(x) - v(x) |.$$

Also let

$$V = \{u, v \in \mathcal{W} : \rho(u, v) < \infty \text{ and } u, v \text{ are continuous}\}.$$

The subset V is introduced to insure the property of completeness. We note that, in general, V will not be uniquely defined depending on our choice of r (per unit runout cost) and θ (per unit outdate cost). We deal with one specific subset V of $\mathcal{W}$.

Denardo assumed that his space V was the space of bounded functions on R. It was recognized by Porteus [6] that this assumption was really not necessary and that the results of the theory apply to the more general classes of functions indicated above.

For any ordering policy define the mapping H_y on V as

$$H_y v(x) = L(x, y(x)) + \alpha \int_0^\infty v(y(x) - (t-x)^+) f(t)\, dt$$

and

$$Av(x) = \inf_{y \geq 0} \left\{ L(x, y) + \alpha \int_0^\infty v(y - (t-x)^+) f(t)\, dt \right\}.$$

For the infinite horizon case we assume $0 < \alpha < 1$. We then have the following result:

THEOREM 3: The mappings H_y and A are both positive contractions on V. In addition both of these mappings satisfy the monotonicity property, that is, if $u(x) \geq v(x)$ (for all $x \in X$) then $H_y u(x) \geq H_y v(x)$, and similarly for A. Finally there exists an optimal stationary policy $y^*(x)$ which is the solution to the equation

$$u^*(x) = L(x, y^*(x)) + \alpha \int_0^\infty u^*(y^*(x) - (t-x)^+) f(t)\, dt,$$

where u^* is the unique fixed point of A. The proof is obvious from Denardo's and Porteus' work.

The term $u^*(x)$ has the interpretation of being the minimum expected cost, given we are in state x and infinitely many periods remain in the horizon. To see that the infinite horizon problem arises in a natural way as the limit of the finite horizon problem as the length of the horizon gets large, let $v_0 \equiv 0$ $v_n = A^n v_0$. Then it follows that v_n is exactly C_n of Theorem 2 and the function

$$\left\{ L(x, y) + \alpha \int_0^\infty v_n(y - (t-x)^+) f(t)\, dt \right\}$$

will be convex in y. Hence the sequence of functions $y_n \to y^*$.

The computational schemes developed employ precisely this idea that the infinite horizon problem is in a sense the limit of the finite horizon problem. Algorithms for successive approximations are considered in Denardo and Porteus.

REFERENCES

[1] Bessler, S. and A. F. Veinott, Jr., "Optimal Policy for a Dynamic Multi-Echelon Inventory Model," Nav. Res. Log. Quart. *13*, 355–389 (1966).

[2] Bulinskaya, E., "Some Results Concerning Optimum Inventory Policies," Theory of Probability and Its Applications, *9*, 165–177 (1964).

[3] Denardo, E., "Contraction Mappings in the Theory Underlying Dynamic Programming," SIAM Review vol. 9 (1967).

[4] Jennings, J. B., "Inventory Control in Regional Blood Banking Systems," Technical Rept. No. 53, Operations Research Center, MIT, Cambridge, Mass. (1970).

[5] Pierskalla, W. P. and C. Roach, "Optimal Issuing Policies for Perishable Inventory," Technical Rept. No. CP70007, Northwestern University (1971).

[6] Porteus, E., "Some Bounds on Infinite Horizon Discounted Sequential Decision Processes," Rand Corporation Technical Rept. No. P4422 (1970).

[7] Van Zyl, G. J. J., "Inventory Control for Perishable Commodities," unpublished Ph.D. dissertation, University of North Carolina (1964).

[8] Veinott, A. F., "Optimal Policy for Multi-product, Dynamic Non-stationary Inventory Problem," Management Science *12*, 206–222 (1965).

[9] Veinott, A. F., "Optimal Ordering, Issuing, and Disposal of Inventory with Known Demand," Unpublished Ph.D. dissertation, Columbia University (1960).

A STOCK REDISTRIBUTION MODEL

Alan J. Kaplan

AMC Inventory Research Office
Institute of Logistics Research
U.S. Army Logistics Management Center

ABSTRACT

Multi-depot supply systems are subject to stock distribution imbalances; i.e., the fraction of total system stock located at a depot may be too small to support the fraction of system demand expected to be placed on it. In the supply system of concern, a customer is always satisfied if there is stock anywhere in the system. Stock redistributions to correct imbalances may reduce both transportation costs and customer waiting times. A model for determining optimum redistribution quantities is formulated, and a practical method of solution for the two depot case is described. Selected numerical illustrations are given.

1. INTRODUCTION

Multi-depot supply systems are subject to stock distribution imbalances; i.e., the fraction of total system stock located at a depot may be too small to support the fraction of system demand expected to be placed on it. The imbalance can arise because of shifts in demand patterns, or simply because of stochastic drawdowns on depot stocks which differ significantly from mean drawdown rates. In the supply system of concern to us, if an imbalance reaches the point where a customer cannot be satisfied from a local depot while there is stock available elsewhere, a direct shipment from a more distant depot will be made.

Inter-depot transfers to correct stock imbalances can reduce customers' waiting times by moving stock closer to them. As Berkeley and Griswold show [3], transfers can also reduce total transportation costs. Transportation rates per ton mile are greatly reduced if a (rail) carload can be moved at one time. Rarely can carload shipments be utilized in direct customer support, but if a catalog of items is reviewed periodically, carload shipments for inter-depot transfer can be assembled.

Determination of how much to transfer requires a calculation of probabilities which is quite complex. A method of solution is presented here for the two depot case and some sample results are given. Details on how this solution was implemented may be found in [5].

The inter-depot transfer problem has been considered previously in the literature [1, 2, 4]. This body of work concerns a supply system in which a customer can only be satisfied from the closest depot. However, Hadley and Whitin's paper may be of interest to readers interested in extending the results discussed here by use of heuristics.

2. THE MODEL

A time horizon is defined by the expected time until new stock will be received by the depots. Stock for all depots is purchased on one contract and arrives everywhere at approximately the same time; thus the total assets and their initial distribution is fixed.

The impact of a transfer of w units of stock on costs incurred over the time horizon are calculated as:

$$I(w) = C_1(w) - C_2 \times E_1(w) + C_2 \times E_2(w) \tag{2.1}$$

where

$I(w)$ – impact

$C_1(w)$ – transfer cost for w units

C_2 – extra cost per unit if customer is satisfied from distant depot. C_2 may include a penalty cost for delay as well as a transportation cost differential.

$E_1(w)$ – expected number of transferred units which will actually be used by receiving depot during the time horizon.

$E_2(w)$ – expected number of transferred units which will have to be used during time horizon to satisfy customers of shipping depot.

Stock is issued on a first come, first served basis, so that if, due to forecast error, the shipping depot runs out of stock before the receiving depot, some of the transferred assets will have to be routed back to customers of the original shipping depot. This accounts for a positive value for $E_2(w)$.

3. ASSUMPTIONS

Assumptions made are:

(1) Impact of a transfer on costs incurred after the time horizon are zero.

(2) Inter-depot transfer time is 0.

(3) There will be no other transfer made during the time horizon.

(4) Demand during the time horizon on each depot is Compound Poisson with known parameters.

Once a solution is obtained, there are various ways of adjusting it to allow approximately for inter-depot transfer time [5], [6]. If it is assumed that transfers can only be made periodically, assumption (3) can undoubtedly be eliminated by use of dynamic programming at a great increase in complexity. Otherwise, a limit on the number of months supply transferred should be utilized [5].

4. FINDING THE OPTIMUM

Let the transferred assets be numbered for conceptual purposes from 1 to w as depicted in Figure 1 below. In the Figure A_1 and A_2 denote assets before transfer; of receiving depot and shipping depot, respectively.

Let the shipping depot and receiving depot issue (conceptually) from the bottom of their asset stocks upward. Thus, the receiving depot issues its original assets, then the transferred assets in the sequence 1, 2, 3, ... The shipping depot issues its remaining assets $(A_2 - w)$, and then transferred assets are back shipped, if necessary, to satisfy shipping depot customers. These assets are back shipped in the sequence w, $w-1$, . . ., where "w" is the last transferred asset touched by the receiving depot, but the first called upon by the shipping depot and so on.

Define

$P_1(k)$ – probability that transferred asset k is issued to customer of receiving depot.

$P_2(k)$ – probability that transferred asset k is issued to customer of shipping depot.

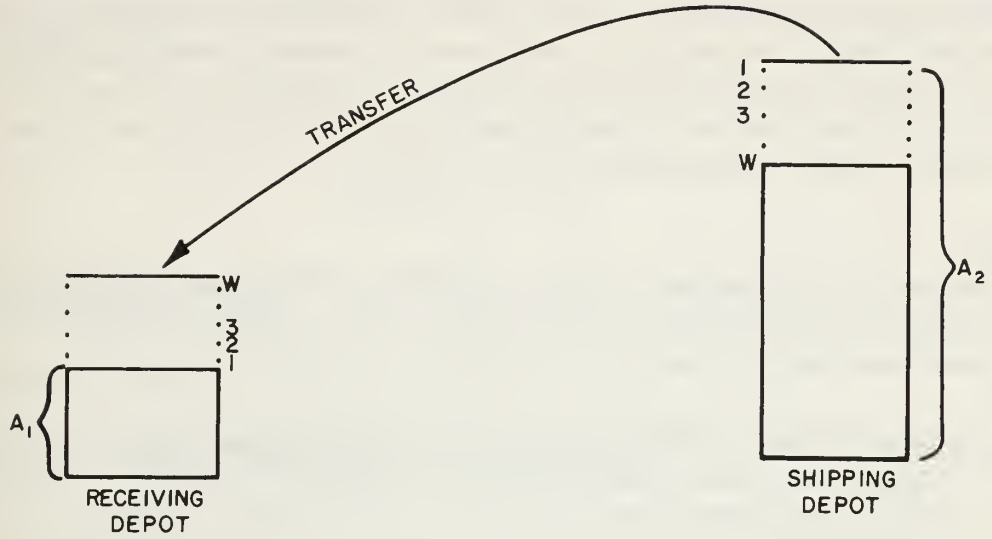


FIGURE 1.

Given the conceptual framework described, $P_1(k)$, $P_2(k)$ are independent of w for $w \geq k$. This is easily illustrated by reference to Figure 1. For example, the condition for use of transfer asset 3 by a receiving depot customer is that:

$$\text{demand on receiving depot} = A_1 + 3$$

before

$$\text{demand on shipping depot} = (A_2 - w) + (w - 2) = A_2 - 2.$$

$P_1(k)$ and $P_2(k)$ are of course not independent of the original asset positions.

By definition of expected value,

$$E_i(w) = \sum_{k=1}^{w} P_i(k) \quad i = 1, 2. \tag{4.1}$$

Hence, if $P_1(k)$, $P_2(k)$ can be calculated for all k, Equation (2.1) can be evaluated for all w and an optimum w found. Finding a w which will minimize (2.1) is made easier if $C_1(w)$ is convex. In that case, $I(w)$ is convex since

$$I(w+1) - I(w) = C_1(w+1) - C_1(w) - C_2 \times [P_1(w+1) - P_2(w+1)]. \tag{4.2}$$

By definition $P_1(w+1) - P_2(w+1)$ cannot increase for increasing w.

5. MATHEMATICS

This section is concerned with the derivation of expressions which make calculation of $P_1(k)$ and $P_2(k)$ straightforward. This derivation is made complex by one consideration—it is quite possible that customers of both the shipping and receiving depots could use the kth redistributed asset sometime during the time horizon. In this case, in accordance with the first come, first served policy, use will be determined by where it is needed first.

The necessity of considering this dynamic aspect of need led to the division of the time horizon into n periods; if depot A's customers need the kth asset in an earlier period than depot B's customers, the asset is credited as being used in support of depot A's customers and vice versa. Note in this the distinction between need and use. Customers of both depots may need the kth asset but only one customer can use it, and it is use to which $P_1(k)$ and $P_2(k)$ pertain.

Notation

We will refer to the customers of a depot frequently in the following. For convenience, the phrase "area j" will be used to denote "the customers of depot j." Also, since we will always be referring to the kth redistributed asset, reference to k is omitted from the notation.

j — index: $j=1$ refers to receiving depot; $j=2$ refers to shipping depot.
n — number of periods into which time horizon is divided.
i — index denoting a period.
U_{ij} — probability that kth (redistributed) asset will be used by area j by the end of period i.
U'_{ij} — probability that use will be during period i.
N_{ij} — probability that kth (redistributed) asset is needed by area j by the end of period i.
N'_{ij} — probability that need will first arise during period i.
$Z1$ — intermediate result equal to $\sum_{j=1}^{2} N'_{ij}$; appropriate i will be clear from context.
$Z2$ — intermediate result equal to $\prod_{j=1}^{2} N'_{ij}$.
E_{ij} — denotes the event that kth redistributed asset is needed by area j before period i.
E'_{ij} — event that need first arises during period i.
E''_{ij} — event that need does not arise until at least after period i.
A_j — initial assets of depot j, before redistribution.
d_{ij} — total demand by area j during periods 1 thru i.
$Pr(\ )$ — denotes probability of event in parenthesis.

Basic Equations

By definition

$$P_j(k) = U_{nj} = \sum_{i=1}^{n} U'_{ij}. \tag{5.1}$$

Also by definition

$$\begin{aligned} N'_{ij} &= N_{ij} - N_{i-1,j} && \text{for } i > 1 \\ &= N_{ij} && \text{for } i = 1. \end{aligned} \tag{5.2}$$

The equations for determining N_{i1} and N_{i2} are obvious. If k assets are redistributed, this gives depot 1 (A_1+k) assets, and leaves depot 2 with (A_2-k) assets. If area 1 has demand of at least (A_1+k),

the kth redistributed asset in needed by area 1. If area 2 has demand greater than (A_2-k), the kth asset was needed there. Thus

$$N_{i1}=Pr\,(d_{i1}\geqq A_1+k)$$

(5.3)

$$N_{i2}=Pr\,(d_{i2}>A_2-k).$$

Using (5.1), (5.2), and (5.3), it is clear that the $P_j(k)$ can be evaluated easily if the U'_{ij} can be evaluated as functions of the N_{ij} and N'_{ij}.

Methodology for $n=1$

An approximate expression is justified for U'_{ij} when the time horizon is not divided into periods ($n=1$). In the next section, the approximation is used in solving U'_{ij} for $i>1$. It is shown that as n is increased, any error introduced into the estimate of $P_j(k)$ by this approximation must approach 0.

For depot 1,

$$U'_{11}=N'_{11}\times(1-N'_{12})+N'_{11}\times N'_{12}\times\frac{N'_{11}}{N'_{11}+N'_{12}} \tag{5.4}$$

$$=N'_{11}\times(1-N'_{12})+Z2\times\frac{N'_{11}}{Z1}.$$

The first term on the right gives the joint probability that the kth asset is needed by area 1 during the time horizon (period 1), and not needed by area 2. In this case, of course, the asset will be used by area 1.

In the event that the kth asset is needed by both areas, an event with probability $N'_{11}\times N'_{12}$, it cannot be precisely determined by which area the asset will be used. As an approximation, it is assumed that the asset will be used by area 1 in the proportion $N'_{11}/(N'_{11}+N'_{12})$. This is the basis for the second term on the right hand side of (5.4).

Equation (5.4) is simplified by algebra to give

$$U'_{11}=N'_{11}-Z2\times\left(1-\frac{N'_{11}}{Z1}\right). \tag{5.5a}$$

By a similar derivation (using symmetry),

$$U'_{12}=N'_{12}-Z2\times\left(1-\frac{N'_{12}}{Z1}\right). \tag{5.5b}$$

As a check note that

$$U'_{11}+U'_{12}=N'_{11}+N'_{12}-Z2\times\left(2-\frac{N'_{11}+N'_{12}}{Z1}\right) \tag{5.6}$$

$$=Z1-2\times Z2+Z2$$

$$=Z1-Z2.$$

Equation (5.6) can be derived directly using this result from elementary probability theory:

$$Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B). \tag{5.7}$$

In this case the events A and B correspond to the need by area 1 for the kth redistributed asset and the need by area 2 for that asset.

Methodology for $n > 1$

Using the theory of conditional probability:

$$\begin{aligned} U'_{i1} = {} & (U'_{i1} \mid E_{i2}) \times Pr(E_{i2}) \\ & + (U'_{i1} \mid E'_{i2}) \times Pr(E'_{i2}) \\ & + (U'_{i1} \mid E''_{i2}) \times Pr(E''_{i2}). \end{aligned} \tag{5.8}$$

Analyzing each term:

(a) if the asset is needed by area 2 before period i, it cannot be used during period i by area 1 in accordance with the first come-first served issue policy.

(b) if the asset is needed during period i by both area 1 and area 2 (second term) the approximation underlying (5.5) is applied with the trivial generalization from period 1 to period i;

(c) if the asset is not needed by area 2 until after i (3rd term), then need by area 1 during period i is synonomous with use there.

Hence,

$$\begin{aligned} & U'_{i1} \mid E_{i2} = 0, \\ & U'_{i1} \mid E'_{i2} = N'_{i1} \times \frac{N'_{i1}}{Z1}, \\ & U'_{i1} \mid E''_{i2} = N'_{i1}. \end{aligned} \tag{5.9}$$

Further

$$\begin{aligned} & Pr(E'_{i2}) = N'_{i2} \\ & Pr(E''_{i2}) = 1 - N_{i2}. \end{aligned} \tag{5.10}$$

Substituting (5.9) and (5.10) into (5.8) and using symmetry yields

$$\begin{aligned} U'_{i1} &= 0 + N'_{i1} \times \frac{N'_{i1}}{Z1} \times N'_{i2} + N'_{i1} \times (1 - N_{i2}) \\ &= N'_{i1} \times \left(\frac{Z2}{Z1} + 1 - N_{i2} \right) \end{aligned} \tag{5.11}$$

$$U'_{i2} = N'_{i2} \times \left(\frac{Z2}{Z1} + 1 - N_{i1} \right).$$

It is possible, as a check, to derive directly the sum of Equations (5.11) just as equation (5.6) was derived directly [5].

It was earlier claimed that by increasing n, any error introduced into the estimate of $P_j(k)$ by the approximate nature of (5.5) and its generalization to period i, must approach 0.

Now, the error arises because it is not possible to determine precisely which depot will use an asset in the event that both depots need it in the same period, with probability $N'_{i1} \times N'_{i2}$. Hence, an upper bound for the error introduced is clearly:

$$\text{Max error} < \sum_{i=1}^{n} N'_{i1} \times N'_{i2}. \tag{5.12}$$

We derive a rather crude upper bound for N'_{ij}, as a function of n, and then use it to show (5.12) goes to 0 as $n \to \infty$.

It is assumed that demand is Compound Poisson (section 3). If demand is a Compound Poisson process, it has stationary increments, see Parzen [7]. This means that if $E(d_j)$ represents expected demand from area j during the time horizon,

$$E(d_{ij}) - E(d_{i-1,j}) = E\,\frac{(dj)}{n}\,. \tag{5.13}$$

But,

$$N'_{ij} \leq E(d_{ij}) - E(d_{i,j-1}). \tag{5.14}$$

The right hand side is the expected demand by area j in period i. By definition of expected value, this must be at least equal to the probability of any demand by area j in period i, which in turn is clearly $\geq N'_{ij}$.

Hence,

$$\text{Max error} \leq \sum_{i=1}^{n} \frac{E(d_1)}{n} \times \frac{E(d_2)}{n} = \frac{1}{n}\, E(d_1) \times E(d_2). \tag{5.15}$$

The right hand side goes to 0 as $n \to \infty$

Some experimentation with the model indicated that little precision was added by making $n > 3$.

Integral Approach*

The $P_i(k)$ can be expressed very nicely in integral form. Let

$F_1(t, x)$ – the c.d.f. of the time until area 1 has demand $\geq x$,
T – the length of the time horizon,
$G_2(x;\ t)$ – the c.d.f. of demand on area 2 during time t.

Then using (5.3)

$$P_1(k) = \int_0^T dF_1\,(t; A_1 + k) \cdot G_2\,(A_2 - k; t). \tag{5.16}$$

*This approach was suggested by the referee, who also contributed to the general clarity of the paper.

An analogous equation can be given for $P_2(k)$.

Equation (5.16) does not suggest a general solution technique which is superior to that already described, using (5.1) and (5.11). In the simple case where the demand distribution in both areas is the simple Poisson, it was found that evaluation of (5.16) reduces essentially to summing $A_2 - k$ scaled Gamma cumulative density functions.

6. NUMERICAL RESULTS

The methodological approach described has been programmed assuming demand is negative binomial. Selected results are given in the table below. No adjustment is made for inter-depot transfer time as it is assumed to be 0. Three subperiods ($n=3$) are used in the computations.

Notation is as in section 2 except that $C_1(w)$ is assumed equal to $C_1 \times w$. Also.

I — denotes depot 1
II — denotes depot 2
AMD — denotes average monthly demand
Coef of Var — denotes coefficient of variation of demand over the time horizon
OPT — denotes optimum transfer quantity
Savings — denotes $|\text{I (OPT)}/(C_1 \times \text{OPT})|$ (refer to Equation 2.1)

In reading the table, all inputs for cases 2–8 are the same as for case 1 except for the one input specified in each case.

Case	AMD		Coef of VAR		Time horizon (months)	C_2/C_1	Initial assets		OPT	Savings (%)
	I	II	I	II			I	II		
1	40	40	0.25	0.25	9	1.7	40	400	138	56.3
2		80							42	40.9
3			0.50						119	47.5
4				0.50					130	50.7
5					12				125	51.6
6						1.2			110	16.5
7							80		113	51.7
8								200	40	42.8

7. EXTENSIONS

No serious attempt was made to extend the formal solution of the inter-depot transfer problem as discussed here to the three-depot case. It presents an interesting problem because there is a cascading of decisions; for example, the solution depends on what the search pattern is when one depot runs out

of stock, there is demand on this depot, and the demand can be supplied from either of the two remaining depots.

REFERENCES

[1] Allen, S. G., "Redistribution of Total Stock Over Several User Locations," Nav. Res. Log. Quart. *5*, 51–59 (1958).

[2] Allen, S. G., "A Redistribution Model With Set Up Charge" and "Computation for the Redistribution Model With Set Up Charge," Stanford Research Institute, Menlo Park, Calif., T.M. #17 and #20 (Jan. 1961 and July 1961).

[3] Berkeley, Robert M. and John W. Griswold, *Dynamic Stock Positioning*, Logistics Studies Office, U.S. Army Logistics Management Center, Ft. Lee, Va. CDT–01–08–71–04 (1971).

[4] Hadley, G. and T. M. Whitin, "A Discrete Review Inventory Model With The Inventory Position Restored Each Period," Stanford Research Institute, T.M. #12 (Feb. 1961).

[5] Kaplan, Alan J., *Methodology for Inter-Depot Transfers*, AMC Inventory Research Office, Institute of Logistics Research, U.S. Army Logistics Management Center, AD736847 (1971).

[6] Kaplan, Alan J. and Robert L. Deemer, *Stock Allocation in a Multi-Echelon System*, AMC Inventory Research Office, Institute of Logistics Research, U.S. Army Logistics Management Center, AD727694 (1971).

[7] Parzen, Emanuel, *Stochastic Processes* (Holden-Day, San Francisco, 1962).

THREE WAYS OF OBTAINING THE AVERAGE COST EXPRESSION IN A PROBLEM RELATED TO JOINT REPLENISHMENT INVENTORY CONTROL*

Edward A. Silver

University of Waterloo,
Waterloo, Ontario.
Canada

ABSTRACT

This paper does not present a new result, rather it is meant to illustrate the choice of modelling procedures available to an analyst in a typical inventory control problem. The same "average cost per unit time" expression is developed by three quite different procedures. This variety of approaches, as well as the recounting of the author's chronological efforts to solve the problem, should be of interest to the reader. The specific inventory problem studied is one where the controller of an item is faced with random opportunities for replenishment at a reduced setup cost; the problem is an integral component of the broader problem of inventory control of a group of items whose replenishments are coordinated to reduce the costs of production, procurement, and/or transportation.

INTRODUCTION

The specific problem to be shown in this paper was encountered in the course of basic research being conducted in the inventory control of items that are not treated independently. In such a context, if we consider the behaviour of a specific item, at times it triggers certain replenishment orders and on other occasions it is included in a replenishment order triggered by some other item. An important question can be stated as follows: Given that a specific type of inventory control policy is in use, namely an $(S, c, 0)$ policy,† we wish to determine the average costs (carrying and replenishment) per unit time of a particular item i as a function of the values assigned to its two control variables, S_i and c_i. Under certain assumptions concerning the demand for item i and the frequency of orders triggered by other items we are led to precisely the problem studied in this paper.

In attacking the problem the first approach used was to recognize that the average costs of inventory control of an item per year can be expressed as

$$C(S, c) = \bar{I}(S, c)vr + \bar{N}(S, c)\cdot(\text{average cost per order}), \tag{1}$$

where $C(S, c)$ indicates that the cost is a function of the two control variables S and c.

*This work was supported by the Defence Research Board of Canada Grant number 9740–16 and by Grant number A7417 from the National Research Council.

†This is a special case of a "can-order" or (S, c, s) policy. In an (S, c, s) policy we have a continuous review system. Whenever item i's inventory level hits s_i or lower it triggers an order so as to raise item i's level to S_i. At the same time any other item j with inventory level at-or-below its can-order-point c_j is included in the replenishment. If item j is included, a quantity is ordered sufficient to raise its level to S_j. References on (S, c, s) policies include Balintfy [2], Silver [12], Ignall [8], Curry et al. [4], and Schaack and Silver [11]. s_i is set equal to zero here because in this paper we shall treat the case of unit-sized transactions and instantaneous replenishment.

$\bar{I}$ is the average inventory level, in pieces,
v is the standard unit cost of the item, in \$/piece,
r is the carrying charge in \$/\$/year,

and $\bar{N}$ is the average number of orders involving the item per year. It should be noted that the average cost per order is also influenced by the values of S and c through their influence on the fraction of the orders for the item that are triggered by the item itself (as will be discussed in the next section, the item triggering the replenishment has an extra fixed cost assigned to it). This type of approach is quite common in the analysis of inventory control problems (see, for example, Naddor [9] or Brown [3]).

Having written Eq. (1), the problem boils down to obtaining analytic expressions (as functions of S and c) for each of $\bar{I}$, $\bar{N}$, and "fraction of orders involving an item that are triggered by that item". The author tackled this problem by considering an intermediate variable z, namely the distance below c at which the item is included in an order. This integer random variable can take on any of the values $0, 1, 2, \ldots, c-1, c$. To illustrate, the fraction of orders involving an item that are triggered by the item itself is given by the probability that z takes on the value c. ($z=c$ implies that the inventory dropped to zero, at which time a replenishment was made). The relationship between $\bar{I}$ and the probability distribution of z is far more complicated, as will be illustrated later in the paper. In any event, this approach was used, and from Eq. (1), after considerable labour, an expression for the average cost per unit time was developed as a function of the two controllable variables S and c.

At this stage, the investigator recalled an article written by Friend [5] that might have been related to this problem. It was more than related; it turned out to be a different analysis (using a Markov model) of the exact same problem. In other words, the wheel had been rediscovered. To make matters worse, the results were inconsistent, i.e., one of the wheels was not round.

In an effort to determine where the error lay, a third approach to inventory control analysis was used (see, for example Ackoff and Sasieni [1], or Ross [10]). This renewal process approach uses the fact that, under quite general conditions,

$$\text{Average cost per unit time} = \frac{\text{Average cost per cycle}}{\text{Average duration of a cycle}}.$$

This third attack produced a result consistent with that of Friend. Considerable search finally revealed an algebraic error in the first approach of the author. Once this was corrected, all three procedures produced the same average cost expression.

In the next section we present a more formal definition of the problem together with some notation. This is followed by an exposition of each of the three approaches. The final section summarizes the paper as well as pointing out some potentially beneficial features of the original complicated approach.

PROBLEM STATEMENT

We use a statement (and associated notation) motivated by the joint replenishment setting from which the problem was originally extracted. First, the assumptions and basic notation are spelled out.

Assumptions

1. We consider a single item that is under continuous review (i.e., an ordering decision can be made at any instant of time).

2. Demands for the item are Poisson.*
3. Inventory carrying costs are proportional to the average inventory level.
4. No shortages are allowed.
5. Replenishment is instantaneous.
6. When the item reaches a zero inventory level it triggers a replenishment to raise its level to S.
7. Opportunities to replenish the item at a reduced setup cost occur according to a Poisson process. If the inventory level is above c when such an opportunity arises, no replenishment is made. However, if the level is at or below c, a quantity is ordered sufficient to raise the level to S.

Basic Notation

λ — Rate (per year†) of Poisson process generating (unit size) demands for the item.

μ — Rate (per year) of Poisson process generating opportunities to replenish under a reduced setup cost.

$\rho = \dfrac{\lambda}{\lambda + \mu}$

v — standard unit cost of the item, in \$/piece.

r — carrying charge, in \$/\$/year.

A — major setup cost (often referred to as the header cost [13]), assigned to an item only if it triggers a replenishment, in dollars.

a — minor setup cost (often referred to as the line cost [13]) assigned to each item (including the trigger item) that is involved in a replenishment, in dollars.

S — order-up-to-level, in pieces.

c — can-order-point, in pieces.

Other notation will be introduced as needed during the derivations.

The Problem

Determine the average costs (carrying plus replenishment) per year for given values of λ, μ, v, $r, A, a, S,$ and c.**

FIRST APPROACH: USING THE UNDERSHOOT DISTRIBUTION

Definition of the Undershoot

The undershoot z is the distance below its can-order-point (c) at which the item is included in an order. z is an integer random variable that can take on the values $0, 1, 2, \ldots, c-1, c$.

Motivation for Using the Undershoot as an Intermediate Variable

The basic cost, Eq. (1), can be rewritten as

$$C = \bar{I}vr + \frac{\lambda}{\bar{Q}}(PA + a), \tag{2}$$

*Friend [5] has an interesting discussion of how to use an "equivalent" Poisson distribution in place of a more general actual demand distribution.

† Any other unit time period could be used as long as one was consistent throughout.

** Obviously then the objective is to use this cost expression to determine the (S, c) pair that makes the costs as low as possible. This more general aspect will not be pursued in this paper; the interested reader is referred to the article written by Friend [5].

where $\bar{Q}$ is the average replenishment quantity for the item under consideration, in pieces,
$\bar{I}$ is the average inventory of the item, in pieces, and
P is the probability that an order involving the item under consideration is triggered by the item itself;
all other quantities are as defined earlier.

Now a little reflection reveals that*

$$P = p_z(c). \tag{3}$$

Furthermore,

$$Q = S - c + z, \tag{4}$$

$$\therefore \bar{Q} = S - c + \bar{z}.$$

The quantity $\bar{I}$ is also related to z, but in a far more complicated manner. To derive the relationship, we proceed as follows:

$$\bar{I} = \frac{\bar{D}_a \bar{I}_a + \bar{D}_b \bar{I}_b}{\bar{D}_a + \bar{D}_b}, \tag{5}$$

where $\bar{D}_a$ is the average duration of the "above c" phase of a cycle† ("above c" means that the inventory level is above c),

$\bar{I}_a$ is the average inventory level during "above c" phases,
$\bar{D}_b$ is the average duration of the "at-or-below c" phase of a cycle, and
$\bar{I}_b$ is the average inventory level during "at-or-below c" phases.

In the "above c" phase the item spends an average time of $1/\lambda$ at each of the levels $S, S-1, \ldots, c+2$ and $c+1$. Therefore,

$$\bar{D}_a = \frac{S-c}{\lambda} \tag{6}$$

and

$$\bar{I}_a = \frac{S+c+1}{2}. \tag{7}$$

In a cycle selected at random the probability that the undershoot is k is given by $p_z(k)$. However, in obtaining $\bar{I}_b$ we are not picking a cycle at random; rather, we are sampling at a number of points

* $p_z(k) =$ prob {random variable z takes on the value k}.

† A cycle is the period between two consecutive replenishments of the item (i.e., an old cycle ends and a new one begins each time the inventory level is raised to S).

picked at random in time from portions where we are at or below c. Therefore, the probability that the cycle in which we land has an undershoot k is given by

$$\frac{p_z(k)\,(\overline{D_b \mid z=k})}{\sum_t p_z(t)\,(\overline{D_b \mid z=t})} \quad \text{or} \quad \frac{p_z(k)\,(\overline{D_b \mid z=k})}{\bar{D}_b},$$

where $(\overline{D_b|z=t})$ is the expected duration of the "at-or-below c" phase of a cycle given that the undershoot is equal to t. Therefore,

$$\bar{I}_b = \sum_{k=0}^{c} \frac{(\overline{I_b|z=k})\,p_z(k)\,(\overline{D_b|z=k})}{\bar{D}_b}, \tag{8}$$

where $(\overline{I_b|z=k})$ is the average inventory level in the "at-or-below c" phase of a cycle given that the undershoot is equal to k. It is clear that, if we can determine expressions for $p_z(k)$, $(\overline{I_b|z=k})$ and $(\overline{D_b|z=k})$, then we have a method for finding $\bar{I}_b$, the latter being needed, via Eq. (5), for insertion in the basic cost Eq. (2).

Probability Distribution of the Undershoot

We think of the undershoot as being the consequence of the events generated by two independent Poisson processes. This is illustrated in Figure 1. x_i represents the time between the $(i-1)$st and ith demand after the item's inventory hits the c level. y represents the time from when the item's inventory first hits the c level until an opportunity for a replenishment at a reduced cost occurs. As indicated, the analysis assumes that the x_i's are independent of the y. For a joint replenishment situation, particularly where there are only a few items in the group that is coordinated for replenishment purposes, this is certainly only an approximation. We would expect that with just a few items in the group the inventory levels of the items might tend to lock into a pattern where knowing that a particular item has a low inventory level would change the likelihood of another triggering an order. Tests have verified that this is indeed the case for the situation where there are only a few items in the group. However, relaxation of the assumption of independent Poisson processes is unlikely to produce major quantitative changes (and certainly not qualitative changes) in the behaviour of the total cost expression as a function of the various parameters involved. Moreover, relaxation of the assumption would make the analysis prohibitively complex.

The time from when the item's inventory first hits the c level until it is replenished is the random variable D_b, the duration of the "at-or-below c" phase of a cycle. D_b is given by

$$D_b = \min\left\{\sum_{i=1}^{c} x_i, y\right\}. \tag{9}$$

Now, we can write the following equivalences of events*

$$[z=k] \equiv \left[\sum_{i=1}^{k} x_i < y < \sum_{i=1}^{k+1} x_i\right], \qquad k=0, 1, \ldots, c-1, \tag{10}$$

*$[A] \equiv [B]$ is to be interpreted as event A is equivalent to event B. Also $\sum_{i=1}^{0} x_i$ is defined to be zero.

$$[z=c] \equiv \left[\sum_{i=1}^{c} x < y\right]. \tag{11}$$

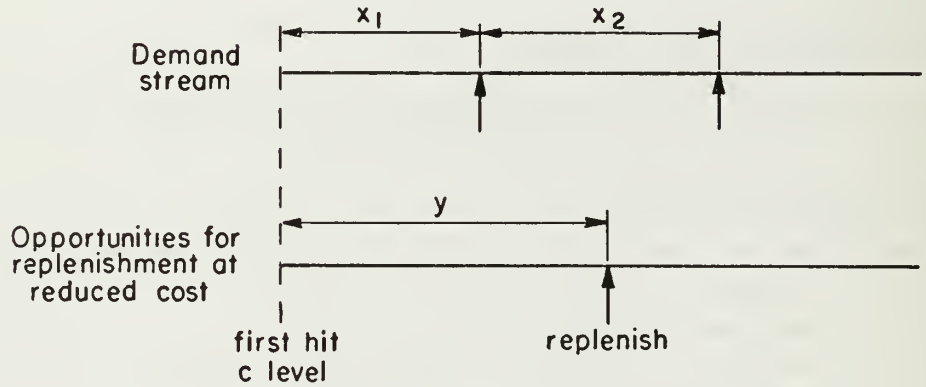
The x_i's are independent draws from an exponential probability density function with parameter λ. Therefore, $\sum_{i=1}^{k} x_i$ has an erlang-type k distribution. Also y is exponentially distributed with parameter μ.

As a consequence of Eq. (10), we have

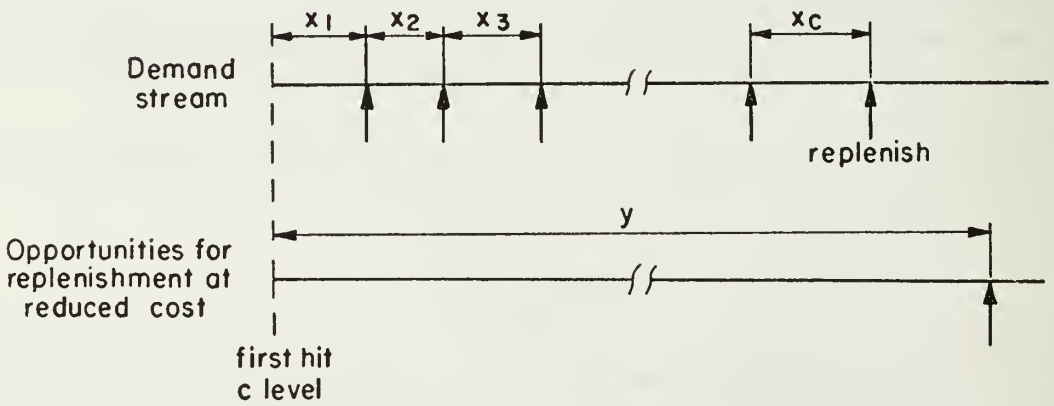
$$p_z(k) = \text{prob}\,[w_k < y < w_k + x_{k+1}], \qquad k=0,1,\ldots,c-1, \tag{12}$$

where w_k has an erlang-type k distribution. The individual density functions of the three random variables are given by †

$$f_w(w_0) = \frac{\lambda^k w_0^{k-1} \exp(-\lambda w_0)}{(k-1)!} \qquad 0 < w_0, \tag{13}$$



(a) Illustration of $z = 1$, $D_b = y$



(b) Illustration of $z = c$, $D_b = \sum_{i=0}^{c} x_i$

FIGURE 1. The undershoot as a consequence of two stochastic processes

† To avoid subscripting a subscript we have suppressed the subscripts on w and x, i.e., $w \equiv w_k$, and $x \equiv x_{k+1}$.

$$f_y(y_0) = \mu \exp(-\mu y_0) \qquad 0 < y_0, \tag{14}$$

and

$$f_x(x_0) = \lambda \exp(-\lambda x_0) \qquad 0 < x_0. \tag{15}$$

Furthermore, the three variables are independent. Therefore, from Eq. (12) we have*

$$p_z(k) = \int_{w_0=0}^{\infty} \int_{y_0=w_0}^{\infty} \int_{x_0=y_0-w_0}^{\infty} f_w(w_0) f_y(y_0) f_x(x_0) \cdot dx_0 dy_0 dw_0. \tag{16}$$

Substitution from Eqs. (13) to (15) and integration gives

$$p_z(k) = \left(\frac{\mu}{\lambda+\mu}\right)\left(\frac{\lambda}{\lambda+\mu}\right)^k \qquad k = 0, 1, 2, \ldots, c-1. \tag{17}$$

Similarly, use of Eq. (11) gives

$$p_z(c) = \left(\frac{\lambda}{\lambda+\mu}\right)^c. \tag{18}$$

Substituting $\rho = \dfrac{\lambda}{\lambda+\mu}$, we have

$$p_z(k) = \begin{cases} (1-\rho)\rho^k & k = 0, 1, 2, \ldots, c-1 \\ \rho^c & k = c, \end{cases} \tag{19}$$

i.e., z has a truncated geometric distribution.

Now, the average value of z is given by

$$\bar{z} = \sum_{k=0}^{c} k p_z(k).$$

The use of Eq. (19) results in

$$\bar{z} = \rho(1-\rho^c)/(1-\rho). \tag{20}$$

Derivation of $(\overline{I_b | z=k})$ and $(\overline{D_b | z=k})$

As in the previous section let

$$w_k = \sum_{i=1}^{k} x_i.$$

*Again, to avoid subscripting a subscript we have suppressed the subscripts on w and x, i.e., $w \equiv w_k$ and $x \equiv x_{k+1}$.

The w_k has an erlang-type k distribution. Furthermore, the conditional probability density function of w_k given $z=k$ is given by*

$$f_{w|z}(w_0|k)=\frac{f_{w,z}(w_0,k)}{p_z(k)}. \tag{21}$$

Case of k$=0, 1, 2, \ldots, c-1$—For this case the numerator of Eq. (21) is the same as the right hand side of Eq. (16) except we do not perform the integration on w_0. Also the denominator is given by Eq. (17). We thus obtain, after some simplification,

$$f_{w|z}(w_0|k)=\frac{(\lambda+\mu)^k}{(k-1)!}\quad w_0{}^{k-1}\exp\left[-(\lambda+\mu)\,w_0\right]\qquad 0<w_0\qquad (k<c). \tag{22}$$

This is still an erlang-type k distribution, but with parameter $(\lambda+\mu)$ instead of λ. In a similar fashion one can show that

$$f_{y|z}(y_0|k)=\frac{(\lambda+\mu)^{k+1}}{k!}\quad y_0^k\exp\left[-(\lambda+\mu)\,y_0\right]\qquad 0<y_0\qquad (k<c),$$

i.e., y also has a conditional distribution that is erlang, but of type $(k+1)$ and with parameter $(\lambda+\mu)$.

A random variable having a distribution that is erlang-type j with parameter γ has an expected value equal to j/γ. Therefore

$$\overline{(w_k\mid z=k)}=k/(\lambda+\mu)$$

and

$$\overline{(y\mid z=k)}=(k+1)/(\lambda+\mu).$$

Hence

$$\overline{(y-w_k\mid z=k)}=1/(\lambda+\mu).$$

But y is the length of the cycle and $(y-w_k)$ is the length of time spent at inventory level $(c-k)$. Therefore, we have proved

$$\overline{(D_b\mid z=k)}=(k+1)/(\lambda+\mu)\qquad k=0, 1, 2, \ldots, c-1, \tag{23}$$

and

$$[\text{expected time at level } (c-k)\mid z=k]=1/(\lambda+\mu)\qquad k=0, 1, 2, \ldots, c-1. \tag{24}$$

---

*$f_{s|t}(s_0|t_0)$ is the probability density function of s evaluated at s_0 given that t takes on the value t_0. Again we have suppressed the subscript on w_k.

By returning to the result of Eq. (22) and recognizing that $w_k = x_1 + x_2 + \ldots + x_k$, it is evident that conditional upon $z = k$, each $x_i (i = 1, 2, \ldots, k)$ has an exponential distribution with parameter $(\lambda + \mu)$. Thus,

$$\overline{(x_i \mid z = k)} = 1/(\lambda + \mu) \qquad i = 1, 2, \ldots, k.$$

but, x_i is the time spent at the inventory level $(c - i + 1)$. Therefore, we have

$$[\text{expected time at level } (c - i + 1) \mid z = k] = 1/(\lambda + \mu) \qquad i = 1, 2, \ldots, k. \tag{25}$$

From the results of Eqs. (24) and (25) it is seen that, given that $z = k$, we spend the same average time in each of the inventory states $c, c-1, c-2, \ldots, c-k$.* Therefore,

$$\overline{(I_b \mid z = k)} = c - \frac{k}{2} \qquad k = 0, 1, 2, \ldots, c-1. \tag{26}$$

Case of k = c—This case has to be handled differently because the cycle ends by the item hitting the zero inventory level, rather than by the occurrence of an opportunity to replenish at a reduced setup cost. Here we can show that†

$$f_{w|z}(w_0 \mid c) = \frac{(\lambda + \mu)^c w_0^{c-1}}{(c-1)\,!} \exp\left[-(\lambda + \mu) w_0\right] \qquad 0 < w_0;$$

i.e., the conditional probability density of w_c is erlang-type c with parameter $(\lambda + \mu)$. It is evident that the conditional distribution of each of the x_i's $(i = 1, 2, \ldots, c)$ must again be exponential with parameter $(\lambda + \mu)$. Therefore, we spend the same average time in each of the states $c, c-1, \ldots, 2, 1$ thus giving

$$\overline{(I_b \mid z = c)} = \frac{c+1}{2}, \tag{27}$$

which is functionally slightly different from Eq. (26). Also, in this case the duration of the "at-or-below c phase" is given by w_c. Hence,

$$\overline{(D_b \mid z = c)} = \overline{(w_c \mid z = c)} = c/(\lambda + \mu). \tag{28}$$

Development of the Cost Expression

Substitution of the results of Eqs. (19), (23), (26), (27), and (28) into Eq. (8) gives

$$\bar{D}_b \bar{I}_b = \sum_{k=0}^{c-1} \left(c - \frac{k}{2}\right) \rho^k (1 - \rho)\ (k+1)/(\lambda + \mu) + \left(\frac{c+1}{2}\right) \rho^c c/(\lambda + \mu)$$

*Furthermore, this average time in each state is independent of k.
† Again, the subscript c is omitted from w_c.

or

$$\bar{D}_b\bar{I}_b\,\lambda = \sum_{k=0}^{c-1}\left(c-\frac{k}{2}\right)\rho^{k+1}\,(1-\rho)\,(k+1)+\left(\frac{c+1}{2}\right)c\rho^{c+1}.$$

Considerable algebra* leads to

$$\bar{D}_b\bar{I}_b\,\lambda = \frac{\rho}{1-\rho}\,[c-\rho(1-\rho^c)/(1-\rho)]. \tag{29}$$

Now

$$\bar{D}_b = \sum_{k=0}^{c}\overline{(D_b\,|\,z=k)}\;\rho_z(k).$$

Using the results of Eqs. (19), (23), and (28), one obtains

$$\bar{D}_b\lambda = \sum_{k=0}^{c-1}(k+1)\,\rho^{k+1}\,(1-\rho)+c\rho^{c+1}.$$

This, in turn, simplifies to

$$\bar{D}_b\lambda = \rho(1-\rho^c)/(1-\rho). \tag{30}$$

Substituting Eqs. (6), (7), (29), and (30) into Eq. (5) gives

$$\bar{I} = \frac{(S-c)\,(S+c+1)/2+\rho[c-\rho(1-\rho^c)/(1-\rho)]/(1-\rho)}{S-c+\rho(1-\rho^c)/(1-\rho)} \tag{31}$$

From Eqs. (3) and (19) we have

$$P = \rho^c, \tag{32}$$

and from Eqs. (4) and (20)

$$\bar{Q} = S-c+\rho(1-\rho^c)/(1-\rho). \tag{33}$$

---

*We use

1. $$\sum_{k=0}^{c-1}\rho^k = (1-\rho^c)/(1-\rho);$$

2. $$\sum_{k=0}^{c-1}k\rho^k = \frac{\rho}{(1-\rho)^2}\,[c\rho^c-c\rho^{c-1}+1-\rho^c];$$

and

3. $$\sum_{k=0}^{c-1}k(k-1)\rho^k = \frac{\rho^2}{(1-\rho)^3}\,[2c(c-2)\rho^{c-1}-c(c-1)\rho^{c-2}-(c-2)\cdot(c-1)\rho^c+2].$$

Substitution of Eqs. (31), (32), and (33) into Eq. (2) finally gives us our desired result, namely

$$(34)\quad C=\{S-c+\rho(1-\rho^c)/(1-\rho)\}^{-1}\{(S-c)(S+c+1)vr/2+\rho[c-\rho(1-\rho^c)/(1-\rho)]vr/(1-\rho)+\lambda\rho^c A+\lambda a\}.$$

SECOND APPROACH: USING A MARKOV MODEL

We shall only outline the approach here. The reader interested in details is referred to the article by Friend [5]. A somewhat more general problem has been investigated by Hurter and Kaminsky [7]. Our notation differs somewhat from that of the other authors.

We let the inventory level immediately after a demand transaction and/or replenishment action represent the state of the system. A state transition is the result of an occurrence of one of the two independent Poisson processes (the demand process with rate λ and the "reduced cost replenishment opportunities" process with rate μ). Proceeding exactly as in the derivation of the probability distribution of the undershoot we can show that

prob {a particular occurrence* is a demand} $=\lambda/(\lambda+\mu)=\rho$

prob {a particular occurrence is a special replenishment opportunity} $=1-\rho$

and

$$(35)\qquad \text{(expected time between occurrences)}=1/(\lambda+\mu).$$

The possible states of the system are $S, S-1, S-2 \ldots, c+1, c, c-1, \ldots, 2, 1$. If we are in any of the states $S, S-1, S-2, \ldots, c+1$, a transition to the next lower state can occur (with probability ρ) or back to the state itself (with probability $1-\rho$ because we ignore opportunities to replenish if the state is $c+1$ or higher). However, from any of the states $c, c-1, \ldots, 3, 2$ a transition can occur to the next lower state (with probability ρ) or to state S (with probability $1-\rho$). Finally, the transition from state 1 must be to state S.

If we let $\underline{\pi}=(\pi_1, \pi_2, \ldots, \pi_S)$ represent the steady-state probability vector, then we can solve for $\underline{\pi}$ by using† $\underline{\pi}=\underline{\pi}P$

and

$$\sum_{i=1}^{S}\pi_i=1,$$

where P is the matrix of transition probabilities described above. The result is

$$(36)\qquad \pi_i=\begin{cases}[S-c+\rho(1-\rho^c)/(1-\rho)]^{-1} & i=c+1, c+2, \ldots S-1, S\\ [S-c+\rho(1-\rho^c)/(1-\rho)]^{-1}\rho^{c+1-i} & i=1, 2, \ldots, c-1, c.\end{cases}$$

The average inventory level is**

$$\bar{I}=\sum_{i=1}^{S} i\pi_i.$$

*An occurrence is either a demand or an opportunity to replenish at reduced cost.

†This is a well-known result of Markov processes (see, for example, Howard [6]).

**Strictly speaking π_i is the probability of being in state i immediately after a transition. However, this is also the probability of being in state i at a random point in time because the same average time is spent in each state between transitions.

Use of Eq. (36) leads to precisely the same result as Eq. (31), namely,

$$\bar{I} = \frac{(S-c)(S+c+1)/2 + \rho[c-\rho(1-\rho^c)/(1-\rho)]/(1-\rho)}{S-c+\rho(1-\rho^c)/(1-\rho)} . \tag{37}$$

To get at the expected replenishment costs per year it is convenient to use $1/(\lambda+\mu)$ as unit time (the average time between transitions). Regular (expensive) replenishment takes place if we are in state 1 and a demand occurs. The probability of this joint event is $\pi_1\rho$. Reduced cost replenishment takes place if we are in any of states $c, c-1, \ldots, 2, 1$ and a replenishment opportunity occurs. The probability of this joint event is $(\pi_c+\pi_{c-1}+\ldots+\pi_2+\pi_1)(1-\rho)$. Therefore,

$$\text{(Expected replenishment cost per transition interval)} = ER$$

$$= \pi_1\rho(A+a) + (\pi_c+\pi_{c-1}+\ldots+\pi_2+\pi_1)(1-\rho)a.$$

Using Eq. (36), we obtain

$$ER = [S-c+\rho(1-\rho^c)/(1-\rho)]^{-1}[\rho^{c+1}A+\rho a]. \tag{38}$$

Now, the expected total cost per year are

$$C(S, c) = \bar{I}vr + ER(\lambda+\mu), \tag{39}$$

where v and r were defined earlier.* Substitution of Eqs. (37) and (38) into Eq. (39) gives precisely the same result as obtained in Eq. (34).†

THIRD APPROACH: USING RENEWAL THEORY

As discussed in the introduction, under quite general conditions

$$\text{Average cost per unit time} = \frac{\text{Average cost per cycle}}{\text{Average duration of a cycle}} .$$

In mathematical notation

$$C(S, c) = \frac{\bar{C}}{\bar{D}} , \tag{40}$$

where $\bar{C}$ is the average cost per cycle, in \$ and $\bar{D} = \bar{D}_a + \bar{D}_b =$ average duration of a cycle, in years. We already have an expression for $\bar{D}$, namely,

$$\bar{D} = (S-c)/\lambda + \rho(1-\rho^c)/(1-\rho)\lambda. \tag{41}$$

---

*The units of r are \$/\$/year. The multiplication of ER by $(\lambda+\mu)$ is to convert the second term to a cost on an annual basis.

†Provided we recognize that $\lambda = \rho(\lambda+\mu)$.

Now

(42)

$$\bar{C}=\left[\left(\frac{S+c+1}{2}\right)\left(\frac{S-c}{\lambda}\right)+\frac{c}{(\lambda+\mu)}+\frac{\rho(c-1)}{(\lambda+\mu)}+\frac{\rho^2(c-2)}{(\lambda+\mu)}+\cdots+\frac{\rho^{c-1}(1)}{\lambda+\mu}\right]vr+\rho^c A+a.$$

The whole expression in square brackets represents the average intentory-years per cycle; in it the first term represents the portion of the cycle above c*; the rest of the terms represent the other inventory levels, each being the level multiplied by the probability of ever getting into it on a single cycle† and by the average time spent in it given that we get into it $(1/(\lambda+\mu))$.

Substitution of Eqs. (41) and (42) into Eq. (40) again gives us the same average cost expression as obtained in Eq. (34) by the first method.

SUMMARY

In this paper we have shown three distinctly different ways of obtaining an average cost expression for a specific inventory control problem. This illustrates the common phenomenon that the solution to a problem is obtainable by a variety of routes. How to choose among these is partially an art developed through experience.

In the particular problem studied, if one was interested only in the average cost expression as a function of the control variables S and c, then clearly the first approach taken by the author was the worst of three illustrated. The author fell into this trap because of a tendency to think in terms of a replenishment cycle as being split into two time phases—the period where inventory is above the reorder point and the period where it is at or below the reorder point. Simple inventory models are usually developed through this line of reasoning. The third method, actually developed originally as a test of the results obtained by the other two methods, turns out to be the simplest of the three approaches. It should be pointed out, however, that the renewal process approach was shortened because of information that was already available from the first approach. The first approach may have some merit when viewed in a broader context. It provides directly more auxiliary information than the other two approaches. This information could be of use to the analyst. To illustrate, suppose that we were to relax the assumption of instantaneous replenishment. Then, to measure the expected costs of shortages or the expected level of customer service, the probability distribution of the undershoot would be a key quantity. On the other hand, the Markov model, through algebraic manipulation, could also provide this undershoot distribution. Aside from showing three approaches to the same problem, perhaps the primary virtue of this paper is its demonstration of the power of Markov models and renewal theory.

ACKNOWLEDGMENT

The author would like to thank a referee for his several helpful comments.

---

*The average duration above c is $(S-c)/\lambda$ and the average inventory level is $(S+c+1)/2$.

†ρ^{c-i} is the probability that we make it down to at least the level i.

REFERENCES

[1] Ackoff, R. L., and M. W. Sasieni, *Fundamentals of Operations Research* (John Wiley & Sons, New York, 1968), chap. 8.

[2] Balintfy, J. L., "On a Basic Class of Multi-Item Inventory Problems," Management Science *10*, 287–297 (Jan. 1964).

[3] Brown, R. G., *Decision Rules for Inventory Management* (Holt, Rinehart and Winston, New York, 1967), p. 227.

[4] Curry, G. L., R. W. Skeith, and R. G. Harper, "A Multiproduct Dependent Inventory Model," AIIE Transactions *II* 263–267 (Sept. 1970).

[5] Friend, J. K., "Stock Control with Random Opportunities for Replenishment," O. R. Quarterly *11*, 130–136 (Sept. 1961).

[6] Howard, R. A., *Dynamic Programming and Markov Processes* (Technology Press & Wiley, New York, 1960), p. 7.

[7] Hurter, A. P. and F. C. Kaminsky, "Inventory Control with Random and Regular Replenishment," J. Ind. Eng. *XIX*, 380–385 (Aug. 1968).

[8] Ignall, E., "Optimal Continuous Review Policies for Two Product Inventory Systems with Joint Set-up Costs," Management Science *15*, 277–283 (Jan. 1969).

[9] Naddor, E., *Inventory Systems* (John Wiley & Sons, New York, 1966), pp. 12–14.

[10] Ross, S. M. *Applied Probability Models with Optimization Applications* (Holden-Day, San Francisco, 1970), p. 52.

[11] Schaack, J. P. and E. A. Silver, "A Procedure for Selecting the Control Variables of an (S, c, s) Joint Ordering Strategy," Infor, *10*, 154–170 (June 1972).

[12] Silver, E. A., "Some Characteristics of a Special Joint Order Inventory Model," Operations Research *13*, 319–322 (Mar.–Apr. 1965).

[13] *Wholesale IMPACT—Advanced Principles and Implementation Reference Manual*, IBM Technical Publications Department, White Plains, New York (1965), p. 65.

INVENTORY MODELS WITH A MIXTURE OF BACKORDERS AND LOST SALES

Douglas C. Montgomery,
M. S. Bazaraa,
and
Ajit K. Keswani
Georgia Institute of Technology

ABSTRACT

This article presents several single-echelon, single-item, static demand inventory models for situations in which, during the stockout period, a fraction b of the demand is backordered and the remaining fraction $1-b$ is lost forever. Both deterministic and stochastic demand are considered, although the case of stochastic demand is treated heuristically. In each situation, a mathematical model representing the average annual cost of operating the inventory system is developed, and an optimum operating policy derived. At the extremes $b=1$ and $b=0$ the models presented reduce to the usual backorders and lost sales cases, respectively.

INTRODUCTION

The problem of determining economic lot sizes in inventory systems has been treated extensively in recent years. Most of this effort has been concentrated on two general situations regarding the demand process when the system is out of stock; i.e., either all demand during the stockout period is backordered or all demand during the stockout period is lost forever. These two cases result in backorders or lost sales models, respectively. However, in many real inventory systems it is more reasonable to assume that only a fraction, say $b(0 < b < 1)$ of the demand during the stockout period can be backordered, and the remaining fraction $1-b$ is lost. For example, if the inventory item is a spare part some customers whose needs are not critical at the present time can wait for the item to be backordered, while others cannot wait and will fill their demand from another source.

Inventory models which consider a mixture of backorders and lost sales have been suggested and even formulated by several authors [1] [3], but not solved. Moreover, the solution is of interest, as it offers further insight into the nature of inventory processes. It can be shown that making the usual assumptions of all backorders or lost sales when in fact a mixture of the two exists can significantly affect inventory costs.

DETERMINISTIC DEMAND

We consider the single-echelon, single item, static demand case. The following definitions apply throughout this discussion:

D = demand per year
Q = the order quantity
C = unit cost of each item (dollars per unit)
I = inventory carrying cost per year, as a percent of C.
A = fixed ordering cost per inventory cycle (dollars per order)

S = total demand per cycle during the stockout period

π = fixed penalty cost per unit short (dollars per unit)

$\bar{\pi}$ = shortage cost per unit period per backorder (dollars per unit period)

π_0 = profit per unit (dollars per unit)

We assume that the mixture of backorders and lost sales during the stockout period is known and constant. Therefore the total number of demands backordered per inventory cycle is bS, and the total number of demands lost is $(1-b)S$. The inventory geometry for this system is shown in Figure 1. Now the annual revenue received will depend on the length of time for which the system is out of stock, and hence on the operating doctrine. Therefore, both revenue and inventory costs depend upon the operating doctrine, and one would not necessarily obtain the same operating policy from a model which minimizes average annual cost as from a model which maximizes the average annual profit. However, by defining lost sales penalty costs to include lost profits as we do below, either a minimum cost or maximum profit formulation yields identical results.

From the above definitions and a consideration of the inventory geometry we can obtain the average annual cost which is the sum of ordering, carrying, and stockout costs, and which includes lost profit, as

(1)

$$K(Q,S)=\frac{AD}{Q+S(1-b)}+\frac{IC(Q-bS)^2}{2[Q+S(1-b)]}+\frac{\pi SD}{Q+S(1-b)}+\frac{\bar{\pi}bS^2}{2[Q+S(1-b)]}+\frac{\pi_0 SD(1-b)}{Q+S(1-b)}.$$

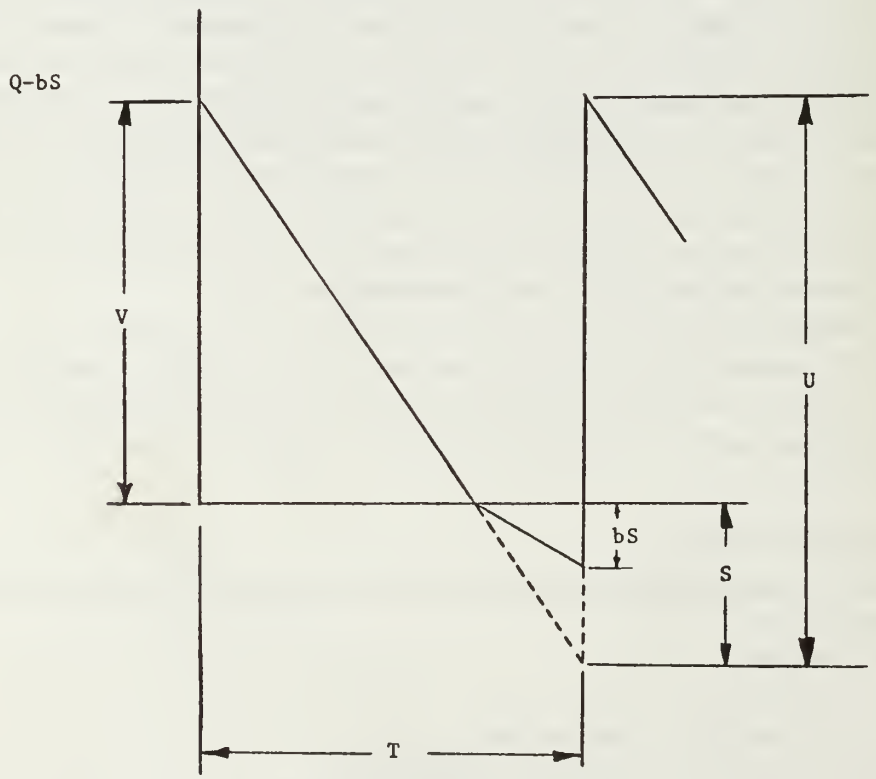


FIGURE 1. A single item deterministic inventory system with the ratio of backorders to demand constant

Note that if $b=1$ or $b=0$, Equation (1) reduces to the usual backorders or lost sales cases, respectively. To find the optimal values of Q and S, say Q^* and S^*, which minimize the average annual cost it is necessary that $\partial K/\partial Q=0$ and $\partial K/\partial S=0$. These conditions result in two simultaneous nonlinear equations in Q and S whose solution is not straightforward. Furthermore, Equation (1) is not convex and we have no guarantee that this will result in a global minimum.

We shall describe a procedure to find the global minimum of K. The following result on nonsingular transformations will be useful:

LEMMA: Let f be a function of $x=(x_1, x_2, \ldots, x_n)$. Define $y=Tx+z$, where $y=(y_1, y_2, \ldots, y_n)$, $z=(z_1, z_2, \ldots, z_n)$, and T is an $(n\times n)$ nonsingular matrix such that $f(x)=f(T^{-1}(y-z))=g(y)$. Then if $g(y^*)\leq g(y)\ \forall\ y\epsilon R^n$, then $f(x^*)=f(T^{-1}(y^*-z))\leq f(x)\ \forall\ x\epsilon R^n$.

PROOF: Since the transformation $y=Tx+z$ is one-to-one, $g(y^*)=f(T^{-1}(y^*-z))\leq g(y)=f(T^{-1}(y-z))\forall y\epsilon R^n$. But $x^*=T^{-1}(y^*-z)$ and $x=T^{-1}(y-z)$. Thus $f(x^*)\leq f(x)\forall x\epsilon R^n$.

Consider the following transformation:

$$U=Q+S(1-b) \tag{2}$$

$$V=Q-bS. \tag{3}$$

Here $T=\begin{bmatrix} 1 & 1-b \\ 1 & -b \end{bmatrix}$, and since $|T|=-1$, we have a nonsingular transformation. We see that U is the total demand during the cycle and $V(V\leq U)$ is the onhand inventory at the beginning of the cycle. Using this transformation Equation (1) becomes

$$\begin{aligned}\hat{K}(U, V) &= \frac{AD}{U}+\frac{ICV^2}{2U}+\frac{\pi D(U-V)}{U}+\frac{\bar{\pi}b(U-V)^2}{2U}+\frac{\pi_0 D(1-b)\ (U-V)}{U} \\ &= U^{-1}[\alpha_1+\alpha_2\ (U-V)+\alpha_3\ (U-V)^2+\alpha_4 V^2],\end{aligned} \tag{4}$$

where $\alpha_1=AD$, $\alpha_2=\pi D+\pi_0 D(1-b)$, $\alpha_3=\bar{\pi}b/2$, and $\alpha_4=IC/2$. Since the transformation (2)(3) is nonsingular, if (U^*, V^*) minimize $\hat{K}$ then by the lemma the inverse transformation yields $(Q^*=bU^*+(1-b)V^*, S^*=U^*-V^*)$ which minimize K. However, $\hat{K}$ is not convex so care must be taken to insure that the global minimum is found.

We minimize $\hat{K}$ in two stages. In the first stage we minimize along the ray $V=\beta U$. Once we find the minimum along each ray of this form the best (minimum cost) ray with the property $\beta\leq 1$ is found. For a given value of β, Equation (4) becomes

$$Z(U)=\hat{K}(U, V=\beta U)=\frac{\alpha_1}{U}+\alpha_2(1-\beta)+U[\alpha_3(1-\beta)^2+\alpha_4\beta^2]. \tag{5}$$

The value of U that minimizes (5) must satisfy $\partial Z/\partial U=0$, which results in

$$\begin{aligned}U^* &= \sqrt{\frac{\alpha_1}{\alpha_3(1-\beta)^2+\alpha_4\beta^2}} \\ &= \sqrt{\frac{2AD}{\bar{\pi}b(1-\beta)^2+IC\beta^2}}\end{aligned} \tag{6}$$

We observe that U^* yields an absolute minimum for (5) since $\partial^2 Z/\partial U^2 = 2\alpha_1 U^{-3} \geq 0$ for $U \geq 0$.

Substituting (6) into (5) yields

$$W(\beta) = \alpha_2[1 - \beta + \sqrt{\alpha_5(1-\beta)^2 + \alpha_6\beta^2}], \tag{7}$$

where $\alpha_5 = 4\alpha_1\ \alpha_3/\alpha_2^2$ and $\alpha_6 = 4\alpha_1\ \alpha_4/\alpha_2^2$. Choosing the best ray involves minimizing W with respect to β. Since $\alpha_2 > 0$ minimizing (7) with respect to β for $\beta \leq 1$ is equivalent to minimizing

$$L(\beta) = -\beta + \sqrt{\alpha_5(1-\beta)^2 + \alpha_6\beta^2} \tag{8}$$

for $\beta \leq 1$. The first derivative $L'(\beta)$ is

$$L'(\beta) = -1 + \frac{-\alpha_5(1-\beta) + \alpha_6\beta}{\sqrt{\alpha_5(1-\beta)^2 + \alpha_6\beta^2}}.$$

It can be verified that $L'(\beta)$ is a monotonically increasing function of β, and furthermore

(a) $L'(0) = -1 - \sqrt{\alpha_5} < 0$, and $L'(1) = -1 + \sqrt{\alpha_6}$,

(b) $L'\left(\frac{\alpha_5}{\alpha_5 + \alpha_6}\right) = -1$,

(c) $L'(\beta) \to -1 + \sqrt{\alpha_5 + \alpha_6}$ as $\beta \to \infty$, and

$L'(\beta) \to -1 - \sqrt{\alpha_5 + \alpha_6}$ as $\beta \to -\infty$.

The graph of $L'(\beta)$ is shown in Figure 2 (a and b).

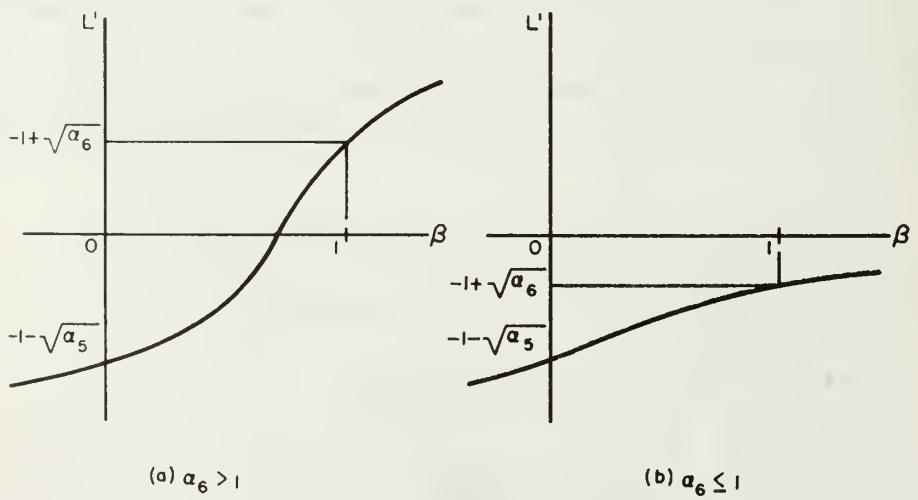


FIGURE 2. Behavior of $L'(\beta)$

We see that if $\alpha_6 \leq 1$ then $L'(1) \leq 0$. Therefore, due to monotonicity, L' is negative over the interval $(-\infty, 1)$. Thus L is decreasing over this interval and the absolute minimum of L is achieved at $\beta^* = 1$. If $\alpha_6 > 1$ then $L'(1) > 0$. Since $L'\left(\frac{\alpha_5}{\alpha_5 + \alpha_6}\right) = -1$ then for some $\beta \epsilon \left(\frac{\alpha_5}{\alpha_5 + \alpha_6}, 1\right)$ we must have $L'(\beta) = 0$. By monotonicity of L' it is clear that this β corresponds to a global minimum. From $L'(\beta) = 0$ we find the optimum β as

$$\beta^* = \frac{\alpha_5}{\alpha_5 + \alpha_6} + \frac{1}{\alpha_5 + \alpha_6}\sqrt{\frac{\alpha_5\alpha_6}{\alpha_5 + \alpha_6 - 1}}. \tag{9}$$

This analysis leads to a simple decision procedure to compute the optimum U^*, V^*. Two cases must be considered:

(i) If $\alpha_6 = \dfrac{2ICA}{D[\pi + \pi_0(1-b)]^2} \leq 1$, then the optimum ray is $\beta^* = 1$. Thus $U^* = V^*$, and from Equation (6)

$$U^* = \sqrt{\frac{2AD}{IC}}.$$

That is, the optimum solution is to allow no backorders or lost sales to occur. The optimum order quantity is $Q^* = U^*$ and $S^* = 0$.

(ii) If $\alpha_6 = \dfrac{2ICA}{D[\pi + \pi_0(1-b)]^2} > 1$ the optimum β^* is given by Equation (9) and the optimum U^* by Equation (6). The optimum value of V is $V^* = \beta^* U^*$. Q^* and S^* can be computed via the inverse transformation.

It is very easy to demonstrate that inventory systems are sensitive to assumptions regarding the nature of demand during the stockout period. For example, suppose that an item has the following characteristics.

$D = 250$ units/year	$b = 0.7$
$C =$ \$10/unit	$\pi =$ \$0.5/unit short
$A =$ \$50/order	$\overline{\pi} =$ \$0.1/unit time per backorder
$I = 0.2$	$\pi_0 =$ \$2/unit lost.

Since $\alpha_6 = 0.661 < 1$ case (i) applies. The optimum order quantity is $Q^* = 112$ and the minimum average annual cost is \$224. We see that it is cheaper to order 112 units every cycle and set the reorder point such that no demands occur when the system out is out of stock then to operate with a mixture of backorders and lost sales. To carry this example further, we observe that even though $b \neq 1$ may inventory managers often *assume* that $b = 1$. That is, they assume that all demands during the stockout period can be backordered even though this is not strictly true. Suppose we do this for the above data. From Equations (2–26) and (2–27) on page 45 of Hadley and Whitin [2], which define the optimal Q and S, say Q' and S', for a backorders model we may compute $Q' = 420$ and $S' = 350$. The true average annual cost that would result from using Q' and S′ may be found by substituting Q' and S' in Equation (1) as \$274. Thus, failure to use the appropriate model has cost management \$271 − 224 = \$50 per year for this single item. The effects of this in a multi-item inventory may be quite substantial.

It is possible to consider many different functional forms for the manner in which the mixture of backorders and lost sales will occur. A constant ratio, as analyzed above, may be satisfactory in many practical situations, but other models may often be easily treated. For example if we assume that the ratio of backorders to demand increases linearly from $p (0 \leq p < 1)$ when the system just goes out of stock to 1 at the end of the cycle just before stock is delivered the average annual cost is identical to (4), except for the term involving time dependent backorder costs, which becomes

$$\frac{\bar{\pi}(4b-1)(U-V)^2}{6U},$$

where $b=(p+1)/2$. We obtain a solution for the optimum system parameters U^* and V^* which is similar to that described above, but with $\alpha_3'=\frac{\bar{\pi}(4b-1)}{6}$ and $b'=(p+1)/3$ substituted for α_3 and b, respectively. Therefore, the analysis for the linearly increasing ratio of backorders to lost sales case is identical to the constant ratio case except for the definition of two constants. We have investigated other situations whose solution is not so simple. However, the identification of the proper functional form may be difficult in practice, and the constant or linear ratio models are probably satisfactory approximations.

STOCHASTIC DEMAND

The models presented in this section are based on the continuous and periodic review models in Chapters 4 and 5 of Hadley and Whitin [2]. We assume that the ratio of backorders to lost sales is constant during the stockout period. Once again, we deal with the single-echelon, single-item, static demand case.

Continuous Review

Consider a reorder point or (Q, r) inventory model. This is often called a continuous review or "transactions reporting" model, because as the stock level reaches r, a quantity Q is ordered and this would require knowledge of the inventory level immediately after each demand. We shall present a heuristic, approximate treatment of the continuous demand case. An analogous development may be given for discrete demand. The following assumptions are required:

1. The unit cost of the item is a constant independent of the order quantity, Q.
2. There is a fixed shortage cost, π, for each unit of demand occurring during the stockout period whether that unit is backordered or lost.
3. There is no time dependent backorder cost, i.e., $\bar{\pi}=0$.
4. The reorder point r, based on the net inventory is positive.
5. There is never more than a single order outstanding.
6. The stockout period during a cycle is small enough to be neglected so that the average number of cycles per year is D/Q, where D is the average annual demand.

If the backorder and lost sales penalty costs are reasonably large (as is often the case), our assumptions and heuristic treatment are reasonable. Further justifications and discussions of these assumptions are given by Hadley and Whitin.

Suppose that the distribution of lead time demand is a continuous density function, say $h(x)$. Now if A is the cost of placing an order, then since the average annual demand is D and since an order

is placed after every Q demands, the average annual cost of placing orders is DA/Q. If the lead time demand is x then the expected demand short at the end of the cycle is given by

$$\eta(r)=\int_r^\infty (x-r)h(x)dx.$$

Thus, the expected number of backorders per cycle are $b\eta(r)$ and the expected demand lost per cycle is $(1-b)\eta(r)$, where b is the ratio of the expected number of backorders to the expected demand short per cycle. The expected net inventory at the beginning of the cycle, assuming that the arrival of an order initiates a cycle is given by

$$Q+r-\mu+(1-b)\eta(r),$$

where μ is the expected lead time demand. Also, the expected net inventory at the end of the cycle is given by

$$r-\mu+(1-b)\eta(r).$$

Therefore, the average annual cost of carrying inventory is

$$IC\left[\frac{Q}{2}+r-\mu\right]+IC(1-b)\eta(r).$$

The annual fixed shortage cost is $D/Q\,\pi\eta(r)$ and the annual lost profit is $D/Q\,\pi_0(1-b)\eta(r)$.

All the components of the average annual variable cost $K(Q, r)$ have been found. $K(Q, r)$ is just he sum of the above components, or

$$K(Q,r)=\frac{AD}{Q}+IC\left(\frac{Q}{2}+r-\mu\right)+\left[IC(1-b)+\frac{\pi D}{Q}+\frac{\pi_0(1-b)D}{Q}\right]\eta(r). \tag{10}$$

Taking the first partial derivatives of (10) with respect to Q and r, equating to zero and solving yields

$$Q=\sqrt{\frac{2D\,[A+\pi\eta(r)+\pi_0\,(1-b)\eta(r)]}{IC}} \tag{11}$$

and

$$H(r)=\frac{QIC}{QIC(1-b)+\pi D+\pi_0 D(1-b)}, \tag{12}$$

where $H(r)$ is the complimentary cumulative of $h(x)$, that is

$$H(r)=\int_r^\infty h(x)dx.$$

Note that (12) is meaningless if $H(r)>1$. Equation (10) is convex, and hence the solution Q^*, r^* obtained from (11) and (12) yields an absolute minimum. The iterative procedure described by Hadley and Whitin [2] on page 171 may be used to solve Equations (11) and (12).

The effect of b on the minimum average annual cost, say K_b^*, may be examined. Since $\bar{\pi}=0$, the stockout cost per unit of demand is $\pi+\pi_0(1-b)$. Thus the stockout cost per unit of demand is minimum when $b=1$ (backorders model) and maximum when $b=0$ (lost sales model). Thus, $K_{b=1}^* < K_{b=0}^*$, and for $0<b<1$, $K_{b=1}^* < K_b^* < K_{b=0}^*$. Therefore, if $b \neq 1$, and if management could influence the nature of the demand process when the system is out of stock, they would be willing to pay some penalty $Z \leq K_b^* - K_{b=1}^*$ to insure that $b=1$.

Periodic Review

Now consider a heuristic, approximate treatment of the periodic review model, or the (R, T) model. This system requires the inventory level to be reviewed at each time interval of length T, and at each review time a sufficient quantity is ordered to bring the inventory position up to R. We shall treat the continuous demand case, where D is the average demand rate per year. A similar development may be given for discrete demand. The following assumptions are necessary:

1. The cost J of making a review is independent of the variables R and T.
2. The unit cost of the item is a constant, independent of the quantity ordered.
3. There is a fixed shortage cost π for each unit of demand occurring during the stockout period whether that unit is backordered or lost.
4. There is no time dependent backorder cost.
5. The backorders are incurred in very small quantities so that when an order arrives, it is almost always sufficient to meet any outstanding backorders.
6. When the procurement lead time is a random variable, it is assumed that the orders are received in the same sequence in which they are placed, and furthermore, lead times for different orders can be treated as independent random variables.

Once again, Hadley and Whitin discuss the practicality of these assumptions. In many real inventory systems the heuristic model will suffice.

The annual ordering and review cost is given by L/T, where $L=A+J$. To compute the inventory carrying cost, the period T will be used as the time between the arrival of two successive orders rather than between the placement of two successive orders. Let $h(x; T)dx$ be the probability that x units are demanded in a time T. Hadley and Whitin discuss the nature of $h(x; T)$ when procurement lead time is either a constant or a random variable. Now the expected number of demands short per review period is

$$\gamma(R)=\int_R^\infty (x-R)h(x; T)dx.$$

Thus $b\gamma(R)$ demands are backordered per cycle and $(1-b)\gamma(R)$ demands are lost per cycle. The expected net inventory at the beginning of the period is

$$R-\mu+(1-b)\gamma(R),$$

and the expected net inventory at the end of the period is

$$R-\mu-DT+(1-b)\gamma(R),$$

where μ is the expected value of lead time demand. Since it is assumed that the portion of time the system is out of stock is small relative to T, the average annual cost of holding inventory is approximately

$$IC\left[R-\mu-\frac{DT}{2}+(1-b)\gamma(R)\right].$$

The average annual fixed shortage cost is $\pi/T\gamma(R)$ and the annual expected lost sales penalty is $\pi_0/T(1-b)\gamma(R)$.

The average annual variable cost can now be written as the sum of the above components, or

$$K(R, T)=\frac{L}{T}+IC\left(R-\mu-\frac{DT}{2}\right)+\left[IC(1-b)+\frac{\pi+\pi_0(1-b)}{T}\right]\gamma(R). \tag{13}$$

For a given T, the value of R which minimizes (13) must satisfy $\partial K/\partial R=0$. If $H(R; T)=\int_R^\infty h(x; T)dx$ is the complimentary cumulative of $h(x; T)$, then the optimum R is the solution to

$$H(R; T)=\frac{ICT}{ICT(1-b)+\pi+\pi_0(1-b)}, \tag{14}$$

if $H(R; T)\leq 1$. For a given T the solution to (14) will yield an absolute minimum for (13). Equation (13) may be optimized with respect to R and T by tabulating the minimum $H(R, T)$ with respect to R alone for various values of T, and then choosing the T which produces a minimum value of $H(R, T)$ in this table. The behavior of the minimum average annual cost with respect to b, say K_b^* is identical to that of the continuous review model, and management would always be willing to pay some penalty to insure that $b=1$, as this would reduce the minimum average annual cost which could be achieved.

CONCLUSION

This article has treated inventory processes in which a fraction b of demand during the stockout period is backordered and the remainder is lost forever. In addition to a deterministic model heuristic, approximate treatments of reorder point and periodic review models for stochastic demand have been briefly considered. It is felt that these models are representative of the real nature of inventory systems, and may be useful for the practical solution of inventory problems.

REFERENCES

[1] Fabrycky, W. J. and Banks, J., *Procurement and Inventory Systems, Theory and Analysis* (Reinhold Publishing Corporation, New York, 1967).

[2] Hadley, G. and Whitin, T. M., *Analysis of Inventory Systems* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1963).

[3] Jelen, F. C. (Editor), *Cost and Optimization Engineering* (McGraw-Hill, Inc., New York, 1970).

THE LOWER BOUNDED AND PARTIAL UPPER BOUNDED DISTRIBUTION MODEL

Harvey M. Wagner

Yale University

ABSTRACT

This paper provides a smaller equivalent bounded variable transportation problem than that in Charnes, Glover, and Klingman [1] for the lower bounded and partial upper bounded distribution model.

I. INTRODUCTION

Charnes, Glover, and Klingman (see Ref. [1]) describe the following model. A total supply of N units is available for shipment. There are m origins, O_i, $i \epsilon I = \{1, 2, \ldots, m\}$, each of which must distribute at least a_i units of the total supply. There are n destinations, D_j, $j \epsilon J = \{1, 2, \ldots, n\}$, each of which demands at least b_j units of the total supply. The objective is to minimize the total shipping cost, where the cost of shipping from origin O_i to destination D_j is denoted by c_{ij}. It is assumed that $N \geq \Sigma\, a_i, N \geq \Sigma\, b_j$, and that all supplies and demands are nonnegative.

The above authors observe that the problem can be stated as an ordinary linear programming model with $m+n+1$ inequality constraints in mn variables. They propose an equivalent bounded-variable transportation problem with $(m+1)$ origins and $(n+1)$ destinations [that is, $(m+n+2)$ equality restrictions] in $(m+1)(n+1)$ nonnegative variables, $(n+m+1)$ of which also are restricted by upper bounds.

This paper suggests an alternative transportation formulation with $(m+1)$ origins and $(n+1)$ destinations that eliminates the need for the upper bounded variables. One computational advantage of the alternative model is that no recourse to an upper-bounded variable algorithm is required; a second advantage is that the size of a basis for this transportation model formulation, namely, $(m+n+1)$ variables, agrees with the size of a basis in the ordinary linear programming version.

II. ALTERNATIVE EQUIVALENT MODEL

The following observations motivate the alternative formulation. If any units above the minimal supply requirements and up to the amount $N - \Sigma\, a_i$ are shipped to a destination D_j, then the cheapest origin available to D_j will ship. Similarly, if any units above the minimal demand requirement and up to the amount $N - \Sigma\, b_j$ are shipped from an origin O_i, then the cheapest destination available to O_i will receive. And if any units are routed from an origin O_i to a destination D_j implying a shipping schedule that exceeds *both* the minimum levels a_i and b_j, then c_{ij} must be the cheapest alternative in the network—clearly such a routing is not utilized if all $c_{ij} > 0$.

Then in lieu of Problem 2 in Ref. [1], we have the model

$$(1) \qquad \text{minimize} \sum_{\substack{i \epsilon I' \\ j \epsilon J'}} c_{ij} x_{ij}$$

subject to

$$\sum_{j\in J'} x_{ij} = a_i \qquad i\in I \tag{2}$$

$$\sum_{i\in I'} x_{ij} = b_j \qquad j\in J, \tag{3}$$

$$\sum_{i\in I'} x_{i,n+1} = N - \sum_{j\in J} b_j = b_{n+1}, \tag{4}$$

$$\sum_{j\in J'} x_{m+1,j} = N - \sum_{i\in I} a_i = a_{m+1}, \tag{5}$$

and

$$x_{ij} \geq 0 \qquad i\in I', j\in J', \tag{6}$$

where $I' = \{1, 2, \ldots, m+1\}$, $J' = \{1, 2, \ldots, n+1\}$, and

$$c_{i,n+1} = \min_{j\in J} c_{ij} \qquad i\in I, \tag{7}$$

$$c_{m+1,j} = \min_{i\in I} c_{ij} \qquad j\in J, \tag{8}$$

$$c_{m+1,n+1} = \min_{i\in I,\, j\in J} (0, c_{ij}). \tag{9}$$

In this formulation, of course, the value of x_{ij} need not represent the *entire* shipment from O_i to D_j, and $x_{m+1,n+1}$ is a dummy shipment unless $c_{m+1,n+1} < 0$.

The above model permits a solution that ships less than the total available supply of N units. If, instead, the solution is contrained to ship the entire available supply of N units, then the only change needed in the formulation is to let

$$c_{m+1,n+1} = \min_{i\in I,\, j\in J} c_{ij}; \tag{10}$$

in this case, $x_{m+1,n+1}$ corresponds to a real shipment.

The other applications in Ref. [1] also carry over immediately to the alternative formulation.

III. PROOF

The equivalence of the alternative model is conveniently demonstrated by recourse to network representations. The original formulation can be written as

$$\text{minimize} \sum_{\substack{i\in I \\ j\in J}} c_{ij}x_{ij}, \tag{11}$$

subject to

$$\sum_{j\in J} x_{ij} - y_{m+1,i} = a_i \qquad i\in I, \tag{12}$$

$$\sum_{i\in I} x_{ij} - y_{j,n+1} = b_j \qquad j\in J, \tag{13}$$

and

$$\sum_{\substack{i\in I\\ j\in J}} x_{ij} + z_{m+1,n+1} = N. \tag{14}$$

To derive a network representation, perform the following linear transformations in order: multiply each equality in (13) by -1 and subtract each equation in (12) from (14); add all the resultant equations to form a redundant equation and multiply this redundant equation by -1. The associated transshipment network for the case $i=j=2$ is shown in Figure 1. This example is sufficiently general to indicate the proof.

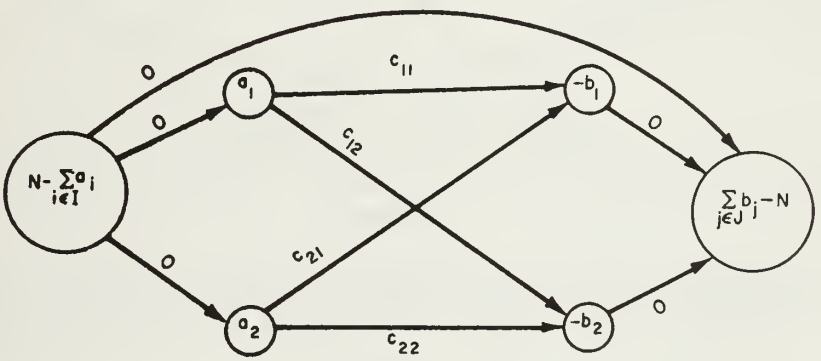


FIGURE 1.

Consider any optimal feasible network flow in Figure 1. If any of the supply $N-\sum_{i\in I} a_i$ is routed to the node with the requirement $-b_1$, then for the cost to be minimal, the arc associated with $\min(c_{11}, c_{21})$ should be utilized; similarly, if any of this supply is routed to the node with the requirement $-b_2$, then the arc associated with $\min(c_{12}, c_{22})$ should be employed. And if any of this supply is routed (directly or through intermediate nodes) to the node with the requirement $\sum_{j\in J} b_j - N$, then for cost to be minimal, the flow should be routed along an arc associated with $\min_{i\in I,\, j\in J} (c_{ij}, 0)$. In other words, an optimal flow in Figure 1 can be made to correspond to a feasible flow in Figure 2, where the cost of the associated flow in Figure 2 equals that in Figure 1. By an analogous line of reasoning, the network in Figure 2 can be modified to that in Figure 3, which is the network for model (1) through (9).

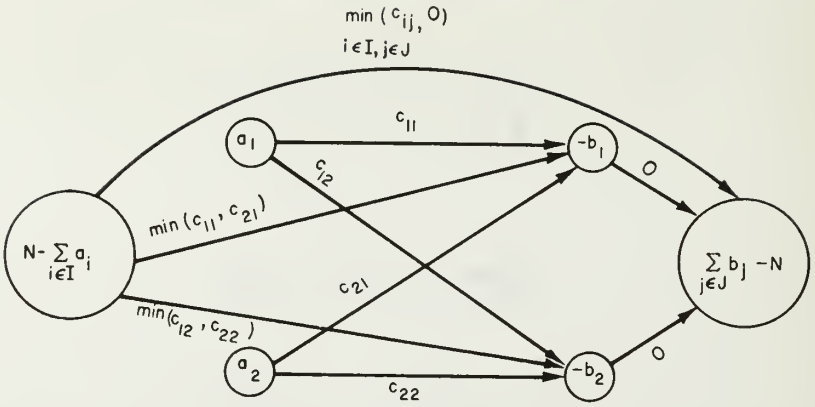


FIGURE 2.

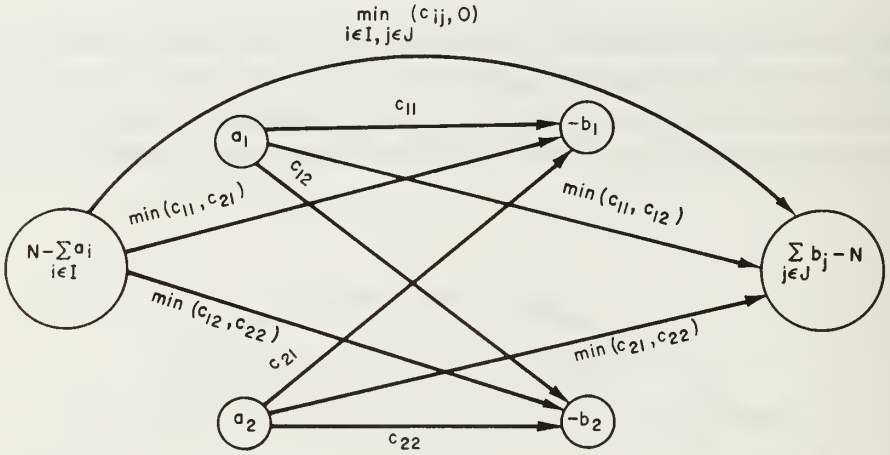


FIGURE 3.

Thus, an optimal flow in Figure 1 can be made to correspond to a feasible flow in Figure 3, where the cost of the associated flow in Figure 3 equals that in Figure 1. To complete the equivalence argument, observe that any flow in Figure 3, and hence an optimal flow in Figure 3, corresponds to a feasible flow in Figure 1 with the same cost.

A more succinct argument follows from the observation that any transshipment network, such as Figure 1, can be transformed to an ordinary transportation problem such as Figure 3, by connecting each source node to each sink node in the original network by an arc whose unit cost in the transformed network equals the minimum cost route from that source to that sink in the original network (see Ref. [2], and pp. 172–173 in Ref. [3]).

REFERENCES

[1] Charnes, A., Fred Glover, and D. Klingman, "The Lower Bounded and Partial Upper Bounded Distribution Model," *Nav. Res. Log. Quart. 18*, 277–281 (1971).

[2] Orden, Alex, "The Transshipment Problem," *Management Science 2*, 276–285 (1956).

[3] Wagner, Harvey M., *Principles of Operations Research With Applications to Managerial Decisions* (Prentice-Hall, 1969).

PEAK RATE OF OCCURRENCE OF A POISSON PROCESS

Khursheed Alam*

Clemson University

ABSTRACT

Suppose that a nonhomogeneous Poisson process is observed for a length of time T, say Let $\lambda(t)$ denote the mean value function of the process. It is assumed that $\lambda(t)$ is first increasing then decreasing inside the interval $(0, T)$ with peak at $t=t_0$, say. Three methods are given for estimating t_0. One of these methods is nonparametric, and the other two methods are based on the standard regression technique and the maximum likelihood principle.

The given result has application in a problem of determining the azimuth of a target from the radar-impulse data. The time series of incoming signals may be approximated by the occurrence of a nonhomogeneous Poisson process with mean value function $\lambda(t)$. The azimuth of the target is reasonably determined from the direction of the axis of the radar beam at the instant t_0, corresponding to the peak value of $\lambda(t)$.

1. INTRODUCTION

The problem of detecting underlying pattern or relationship from a time series data arises in many practical situations. A simple model for the data for such a problem is often of the form

$$X_t=\lambda(t)+\epsilon_t, \qquad t=1, \ldots, n,$$

where $\{\epsilon_t\}$ is a sequence of independent error terms with zero mean, and $\lambda(t)$ is the mean value function. While generally it may be required to estimate $\lambda(t)$, sometimes the primary concern is about any particular characteristic of the mean value function, such as the value of t for which $\lambda(t)$ is maximized or minimized. In this context, a particular model which has been studied by several authors (see Bhattacharya and Johnson [1], Chernoff and Zacks [2], Hinkley [4], [6], Hinkley and Hinkley [7] and Page [8], [9]) is given by

$$\lambda(t = \begin{cases}\lambda_0(t), t=1, \ldots, \tau \\ \lambda_1(t), t=\tau+1, \ldots, n,\end{cases}$$

where $\lambda_0(t)$ and $\lambda_1(t)$ are either constants or linear functions. The primary interest concerns inference about the change point $t=\tau$ in the mean value function.

In this paper we consider a similar problem dealing with the time series data representing a nonhomogeneous Poisson process. The model considered in this paper is appropriate for the data representing impulses received at a radar station from the impact of the radar beam with a distant object, and on account of "noise." The impulses are assumed to represent the occurrence of a nonhomogeneous Poisson process with the mean value function $\lambda(t)$. From the intervals between the successive impulses we estimate the value of t maximizing $\lambda(t)$, that is, the value of t, equal to t_0, say corresponding to the

*The author's work was supported in part by the Office of Naval Research under Contract N00014-71-0339-0002, Task NR 042-271.

peak occurrence rate. The direction of the axis of the radar beam at the instant t_0 provides the most probable value of the azimuth of the object.

The problem of estimating the peak of the mean value function is slightly related to the problem of estimating the mode of a density function. See, for example, Parzen [10]. However, different methods of estimation are involved in the two problems.

MAIN RESULTS

Let a nonhomogeneous Poisson process with mean value function $\lambda(t)$ be observed for a length of time T, and let $\lambda(t)$ be first increasing then decreasing continuously inside $(0, T)$ with peak at $t=t_0$, say. We consider the problem of estimating t_0. We shall assume that the process has occurred at $t=0$ for the reason that for many processes, excluding the Poisson process, the length of time measured from an arbitrary point to the first occurrence does not have the same distribution as the interval between successive occurrences.

First we apply the standard regression method. Let m be an integer such that there is no appreciable change, presumably, in the value of $\lambda(t)$ during the time of m occurrences of the process. It is desirable (see Cox and Lewis [3], section 3.2) that $m \geq 4$. Let Y_1 be the observed time from the start to the mth occurence, Y_2 the time from the mth to the $2m$th occurrence, and so on. If the total number of occurrences during the time T is not a multiple of m, we omit a few occurrences. Then we have a series of intervals $Y_1, \ldots, Y_r$, say. The Y_i's are independently distributed. If the process is a simple Poisson process, that is, if $\lambda(t)$ is a constant function equal to λ, say, then $2\lambda Y_i$ has the Chi-squared distribution with $2m$ degrees of freedom, for each $i=1, \ldots, r$. Therefore,

$$E \log Y_i = \psi(m) - \log \lambda, \quad \operatorname{Var}(\log Y_i = \psi'(m), \tag{2.1}$$

where $\psi(x) = \partial \log \Gamma(x)/\partial x$ denotes the digamma function, and $\psi'(x) = \partial \psi(x)/\partial x$. It is known that the log Chi-squared distribution is asymptotically normal for large value of the degree of freedom of the Chi-square.

Now, let

$$\log \lambda(t) = \alpha + \beta t + \gamma t^2, \ 0 \leq t \leq T, \tag{2.2}$$

where $\beta > 0$ and $\gamma < 0$. The parabolic function $\log \lambda(t)$ is maximized for $t = t_0 = -\dfrac{\beta}{2\lambda}$. We shall assume that $0 \leq t_0 \leq T$.

Let z_i denote the center of the time period Y_i, and let

$$\log \lambda_i = \alpha + \beta z_i + \gamma z_i^2 \tag{2.3}$$

$i=1, \ldots, r$. We make the approximation that $\lambda(t)$ is effectively a constant λ_i within the interval Y_i. We have the linear model

$$EV_i = -\alpha^* - \beta z_i - \gamma z_i^2$$

$$\operatorname{Var}(V_i) = \psi'(m),$$

where $V_i = \log Y_i$ and $\alpha^* = \alpha - \psi(m)$. Let $V = V(V_1, \ldots, V_r)$, $\theta = (\alpha^*, \beta, \gamma)$ and let $A = (a_{ij})$ denote

the matrix of the independent variables, given by $a_{ij} = z_i^{j-1}$ for $i = 1, \ldots, r$ and $j = 1, 2, 3$. The standard least-squares estimate of θ and the covariance of the estimate are given by

$$\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3) \tag{2.4}$$

$$= -VA(A'A)^{-1}$$

and

$$\text{Cov}\,(\hat{\theta}) = \psi'(m) E_A (A'A)^{-1}, \tag{2.5}$$

where A' denotes the transpose of A, and E_A denotes expectation with respect to the distribution of the z_i's. The conditional distribution of $\hat{\theta}$ given $z_1, \ldots, z_r$ is approximately multivariate normal for moderately large values of m.

Let $\hat{t} = -\hat{\theta}_2/(2\hat{\theta}_3)$. An estimate of t_0 is given by

$$\hat{t}_0 = \begin{cases} \hat{t} \text{ for } 0 \leq \hat{t} \leq T \\ 0 \text{ for } \hat{t} < 0 \\ T \text{ for } \hat{t} > T. \end{cases} \tag{2.6}$$

The distribution of $\hat{t}$ is approximated by the distribution of the ratio of two normally distributed random variables. The latter distribution has been investigated by Hinkley [5]. The distribution of the ratio may be approximated by the normal distribution as follows: Let (X, Y) be distributed according to the bivariate normal distribution with mean (μ, ν) and covariance $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. Let $Z = X/Y$. Suppose that μ and ν are large. Applying the δ-method we get

$$Z = X/Y$$

$$= (\mu + \delta_1)/(\nu + \delta_2)$$

$$\sim \frac{\mu}{\nu}\left(1 + \frac{\delta_1}{\mu}\right)\left(1 + \frac{\delta_2}{\nu}\right)^{-1}$$

$$\sim \frac{\mu}{\nu}\left(1 + \frac{\delta_1}{\mu} - \frac{\delta_2}{\nu}\right).$$

Therefore, Z is approximately normally distributed with mean μ/ν and variance equal to

$$\frac{\mu^2}{\nu^2}\left(1 + \frac{1}{\mu^2} + \frac{1}{\nu^2} - \frac{2\rho}{\mu\nu}\right).$$

A few comments are in order on the assumptions made in (2.3) which are rather oversimplifying. If $\lambda(t)$ is a slowly varying function of t or if $\lambda(t)$ is large, in which case the Y_i's would tend to be small,

then it is reasonable to assume that $\lambda(t)$ is approximately constant equal to λ_i, say, for each time period Y_i. This assumption is standard, see for example, Cox and Lewis ([3], section 3.2). Moreover, for verification the m inter arrival time periods which make up Y_i, being considered as m independent observations from m exponential distribution with different scale parameters, may be tested for the equality of the parameters.

The validity of the regression analysis may be questionable as Z_i depends on Y_i. However, as $Z_i = Y_1 + \ldots + Y_{i-1} + \frac{1}{2} Y_i$ we see that the dependence decreases as i increases.

For the adequacy of the relation (2.3) we might consider including a third degree term in the right hand side. However, in view of the several approximations already involved it appears that the inclusion of the additional term would not improve the approximation.

Next we apply the maximum likelihood principle. Suppose that the process occurs at times t_1, t_2, . . ., t_n during the interval $(0, T)$. Let $\xi(T) = \exp\left(-\int_0^T \lambda(t)dt\right)$. The likelihood is given by

$$L = \xi(T) \prod_{i=1}^{n} \lambda(t_i).$$

For the mean value function given by (2.2), we have

$$L = \xi(T) \exp\left(n\alpha + \beta \sum_{i=1}^{n} t_i + \gamma \sum_{i=1}^{n} t_i^2\right).$$

Given β and γ, a sufficient statistic for α is n. Therefore, for inference about β and γ we consider the conditional distribution of the observations, given n. The log likelihood of the conditional distribution is given by

(2.7)

$$l(\beta, \gamma) = -n \log\left(\int_0^T e^{\beta t + \gamma t^2} dt\right) + \beta \sum_{i=1}^{n} t_i + \gamma \sum_{i=1}^{n} t_i^2 + \log n!$$

$$= n\left(\frac{\beta^2}{4\gamma} + \frac{1}{2} \log\left(-\frac{\gamma}{\pi}\right)\right) - n \log\left(\Phi\left(\sqrt{-2\gamma}\left(T + \frac{\beta}{2\gamma}\right)\right) - \Phi\left(\frac{\beta}{2\gamma}\sqrt{-2\gamma}\right)\right) + \beta \sum_{i=1}^{n} t_i + \gamma \sum_{i=1}^{n} t_i^2 + \log n!,$$

where $\Phi(x)$ denotes the standard normal distribution function. The right hand side of (2.7) modified by an additive constant is obtained also by maximizing $\log L$ with respect to α. The estimates of β and γ are obtained from the solution of the likelihood equations $\partial l(\beta, \gamma)/\partial\beta = 0$ and $\partial l(\beta, \gamma)/\partial\gamma = 0$.

Let

$$f(t; \beta) = e^{\beta t + \gamma t^2} / \int_0^T e^{\beta t + \gamma t^2} dt, \qquad 0 \leq t \leq T.$$

$f(t; \beta)$ represents a probability density function with monotone likelihood ratio property. That is,

$$f(t; \beta) f(t'; \beta') - f(t; \beta') f(t'; \beta) \geq 0,$$

for $t' > t$ and $\beta' > \beta$. From Lemma 2.1 given at the end of this section it follows that $\eta(\beta) = \int_0^T tf(t; \beta)dt$ is monotone increasing in β. Now

$$\partial l(\beta, \gamma)/\partial\beta = -n\eta(\beta) + \sum_{i=1}^{n} t_i.$$

From the monotonicity property of $\eta(\beta)$ it follows that the likelihood equation $\partial l(\beta, \gamma)/\partial\beta = 0$ gives a unique solution $\beta = \beta_0(\gamma) =$ say. Similarly, from the likelihood equation $\partial l(\beta, \gamma)/\partial\gamma = 0$ we obtain a unique solution $\gamma = \gamma_0(\beta)$, say. The maximum likelihood estimates $\hat{\beta}$ and $\hat{\gamma}$ are obtained from the intersection of the two curves $\beta = \beta_0(\gamma)$ and $\gamma = \gamma_0(\beta)$. The maximum likelihood estimate of t_0 is set equal to $-\hat{\beta}/2\hat{\gamma}$.

The asymptotic variance and covariance of the maximum likelihood estimates are given by the inverse of the information matrix. The elements of the matrix are negative of the second order partial derivatives with respect to β and γ of the right hand side of (2.7).

The method given above is also applicable when the observed time period $(0, T)$ extends to a preassigned number of occurrences of the process, n, say. Then we maximize the likelihood of the conditional distribution given t_n.

Finally, we apply a nonparametric technique. It is applicable to the case in which no assumption is made with regard to the functional form of the mean value function $\lambda(t)$ except that it is first increasing inside the interval $(0, T)$. Let $X_1, \ldots, X_{n+1}$ denote the time intervals between successive occurrences of the given process, and let $Z_1, \ldots, Z_n$ denote the derived sequence of plus and minus signs, such that $Z_i = +$ or $-$ according to whether $X_{i+1} - X_i$ is positive or negative, $i = 1, \ldots, n$. Let

$$\delta(Z_i) = \begin{cases} 1 \text{ if } Z_i \text{ is } - \\ 0 \text{ if } Z_i \text{ is } +, \end{cases}$$

$$R(0) = \sum_{i=1}^{n} (1 - \delta(Z_i)), \quad R(n) = \sum_{i=1}^{n} \delta(Z_i),$$

and

$$R(m) = \sum_{i=1}^{m} \delta(Z_i) + \sum_{i=m+1}^{n} (1 - \delta(Z_i)), \qquad 0 < m < n. \tag{2.8}$$

Let m_0 denote the value of m, maximizing $R(m)$. If $R(m)$ is maximized for several values of m, select one of these randomly, for the value of m_0. Let $\theta(X_i)$ denote the center of the time interval X_i. An estimate of t_0 is given by

$$\tilde{t}_0 = \theta(X_{m_0+1}). \tag{2.9}$$

If $\lambda(t)$ is a constant function then the X_i's are independently and identically distributed. Then

$$P\{\tilde{t}_0 = \theta(X_i)\} = \frac{1}{n+1}, \qquad i = 1, \ldots, n+1 \tag{2.10}$$

approximately, for large n.

In the general case we make the approximation that $\lambda(t)$ is effectively a constant equal to μ_i, say, inside each time interval X_i, $i=1, \ldots, n+1$. Let $\theta_i=\mu_i/\mu_{i+1}$. As $\mu_i X_i$ is distributed according to the exponential distribution, we have

$$E\delta(Z_i)=P\{X_i>X_{i+1}\} \tag{2.11}$$

$$=(1+\theta_i)^{-1},$$

$$\text{Var}\,(\delta(Z_i))=\theta_i/(1+\theta_i)^2, \tag{2.12}$$

and for $j>i$

$$\text{Cov}\,(\delta(Z_i),\delta(Z_j))=\begin{cases}0 & \text{for } j>i+1\\ -\dfrac{\theta_i\theta_{i+1}}{(1+\theta_i)(1+\theta_{i+1})(1+\theta_{i+1}+\theta_i\theta_{i+1})} & \text{for } j=i+1.\end{cases} \tag{2.13}$$

Therefore, from (2.8) we have

$$ER(m)=\sum_{i=1}^{m}(1+\theta_i)^{-1}-\sum_{i=m+1}^{n}(1+\theta_i)^{-1}+(n-m) \tag{2.14}$$

and

$$\text{Var}\,(R(m))=\sum_{i=1}^{n}\theta_i/(1+\theta_i)^2-2\sum_{i=1}^{n-1}\frac{\theta_i\theta_{i+1}}{(1+\theta_i)(1+\theta_{i+1})(1+\theta_{i+1}+\theta_i\theta_{i+1})} \tag{2.15}$$

$$+4\theta_m\theta_{m+1}/(1+\theta_m)(1+\theta_{m+1})(1+\theta_{m+1}+\theta_m\theta_{m+1}).$$

From (2.14) we have that if $\mu_i<\mu_{i+1}$ for $i=1, \ldots, K$ and $\mu_i>\mu_{i+1}$ for $i=K+1, \ldots, n$, and therefore, the mean value function is maximized for a value of t inside the interval X_{k+1} then the expected value of $R(m)$ is maximized for $m=k$. This result justifies the choice for the estimate of t_0, as given by (2.9).

The nonparametric method given above, is applicable to a more general process, such as, the renewal process with "drift," given by

$$X_t=y_t+\lambda(t),$$

where Y_t represents a renewal process and $\lambda(t)$ is the mean value function, representing the drift.

We have outlined three methods of estimating the time corresponding to the peak of the mean value function $\lambda(t)$ of a nonhomogeneous Poisson process, assuming that $\lambda(t)$ is first increasing then decreasing inside the interval $(0, T)$ during which the process has been observed. The rationale of these methods is clear. But, any of the optimal properties of the given estimates are difficult to derive mathematically. However, these properties can be investigated empirically with the application of the Monte-Carlo technique. The results of the investigation will be given in a separate paper.

The lemma which has been applied above, is stated below. The proof is omitted, as the result is known and follows easily from the monotone likelihood ratio property of the assumed distribution.

LEMMA 2.1: If the distribution of a random variable X has monotone likelihood ratio property in θ, and if $g(x)$ is nondecreasing (nonincreasing) in x then $Eg(x)$ is nondecreasing (nonincreasing) in θ.

REFERENCES

[1] Bhattacharya, G. K. and R. A. Johnson, "Non-Parametric Tests for Shift at an Unknown Time Point," Ann. Math. Statist *39*, 1731–1741 (1968).

[2] Chernoff, H. and S. Zacks, "Estimating the Current Mean of a Normal Distribution Subject to Changes in Time," Ann. Math. Statist *35*, 999–1018 (1964).

[3] Cox, D. R. and P. A. W. Lewis, "The Statistical Analysis of Series of Events," *Methuen's Monographs on Applied Probability and Statistics* (1966).

[4] Hinkley, D. V., "Inference About the Intersection in Two-Phase Regression," Biometrika *56*, 495–504 (1969).

[5] Hinkley, D. V., "On the Ratio of Two Correlated Normal Random Variables," Biometrika *56*, 635–639 (1969).

[6] Hinkley, D. V., "Inference About the Change Point of a Sequence of Random Variables," Biometrika *57*, 1–17 (1970).

[7] Hinkley, D. V. and E. A. Hinkley, "Inference About the Change Point in a Sequence of Binomial Random Variables," Biometrika *57*, 477–488 (1970).

[8] Page, E. S., "A Test for a Change in a Parameter Occurring at an Unknown Point," Biometrika *42* 523–527 (1955).

[9] Page, E. S., "On Problems in Which a Change in Parameter Occurs at an Unknown Point," Biometrika *44*, 248–252 (1955).

[10] Parzen, E., "On Estimation of a Probability Density Function and Mode," Ann. Math. Statist *33*, 1065–1076 (1962).

OPTIMAL SEARCH USING TWO NONCONCURRENT SENSORS*

Carl A. Persinger

Daniel H. Wagner, Associates
Paoli, Pennsylvania

ABSTRACT

Search for a stationary target is considered for a situation in which two sensors are available, but cannot be used simultaneously. The cost (in time) of switching from one sensor to the other is ignored, and each sensor is assumed to have perfect discrimination. For a specified class of searches an optimal allocation of search effort is obtained. In the case of a circular normal prior target location distribution, an example is presented in which one sensor is assumed to have a fixed sweep width and the other a stochastic sweep width. An optimal plan is found for this example. This plan produces an allocation of search effort which is expended in a disk by one sensor and in a bounding annulus by the other.

1. INTRODUCTION

In this paper we consider search for a stationary target for a situation in which two sensors are available, but cannot be used simultaneously. It assumed that the sensors have perfect discrimination so that there are no false targets. A search plan which minimizes the mean time to find the target among the plans in the class Ψ defined below is obtained by an application of an optimization technique presented in [8].

We describe the search model in section 2, and in section 3 we present a form of Theorem 3.2 of [8] and apply the theorem to obtain an optimal allocation of effort. For a circular normal target location distribution an example is presented in section 4, and an optimal allocation of search effort is obtained when one sensor has a deterministic sweep width and an exponential local effectiveness function, and the other sensor has a stochastic sweep width with a gamma distribution and an exponential conditional local effectiveness function. The assumption of a stochastic sweep width and an exponential conditional local effectiveness function for one sensor leads to a non-exponential local effectiveness function for that sensor. For this sensor, the conditional probability of detecting the target with a fixed increment of effort, given that the target has not yet been detected, is strictly decreasing. In this sense, the use of this sensor produces a diminishing rate of return.

In the example, the decision whether to use one or both sensors in the search is determined by comparing the mean sweep rates of the two sensors. If both sensors are used, one searches at a point with the stochastic sensor until the mean posterior sweep rate conditioned on the target being at the point is reduced to the sweep rate of the other sensor; subsequent search at the point is conducted using only the other sensor. For this example it is shown that if the a priori mean sweep rate of the stochastic sensor is larger than that of the deterministic sensor, then the optimal allocation of effort requires that effort by the deterministic sensor be placed in a disk and effort by the stochastic sensor in a bounding annulus.

*This research was supported by the Naval Analysis Programs (Code 462), Office of Naval Research under Contract No. N00014-69-C-0435.

2. SEARCH MODEL

We consider search for a stationary target whose a priori target location distribution has a density function f defined on a region R of Euclidean n-dimensional space. The search vehicle is assumed to have two sensors available for the search, but their simultaneous use is prohibited. Examples of such a situation might be a towed sled upon which only one sensor can be mounted at a time, or a two-sensor array where the performance of each sensor is seriously downgraded by concurrent use of the other sensor. *We ignore the cost (in time delay) of switching from one sensor to the other.*

We assume that each sensor has perfect discrimination so that detecting the target is equivalent to finding it. The term "search effort" is used without definition so that the model will fit many search situations. A sensor effort density function defined on R is a function h such that for any Borel set $R' \subset R$,

$$\int_{R'} h(x)\,dx = \text{amount of effort placed in } R' \text{ by the sensor.}$$

We assume that while searching at a point x in R with sensor j, $j=1,2$, the search vehicle maintains a constant speed U_j. The time required to apply effort by sensor j according to the sensor density function h_j is

$$\frac{1}{U_j}\int h_j(x)\,dx.$$

We denote cumulative search time by s.

For search conducted using sensor j, $j=1,2$, we assume there is a local effectiveness function b_j defined on $R \times [0,\infty)$ such that for each x in R and $z \geq 0$, $b_j(x,z)$ is the conditional probability of finding the target with the jth sensor by search time s given that the target is located at x and the jth sensor search effort density is z at x. Denoting differentiation with respect to the last variable by a prime, we assume, for each j,

(1a)
(i) $b_j(x,\cdot)$ is a probability distribution function
(ii) the derivatives $b_j'(x,\cdot)$ exist, are positive, continuous, and strictly decreasing.

Define

$$J_j(x,z) = b_j'(x,z)/(1-b_j(x,z)), \qquad x \epsilon R,\ z \geq 0.$$

We further assume that the derivatives $b_j''(x,\cdot)$ exist and one of the following holds:

(1b)
(iii) $J_1'(x,\cdot) < 0$ and $J_2'(x,\cdot) \leqslant 0$;
(iv) $J_1'(x,\cdot) = J_2'(x,\cdot) \equiv 0$.

We note that functions b_1 and b_2 which satisfy (i), (ii), and (iv) are necessarily exponential so that in this case, for each x in R, we have positive constants $\omega_1(x)$ and $\omega_2(x)$ for which

$$b_j(x,z) = 1-\exp(-\omega_j(x)z), \qquad j=1,2.$$

The constant $\omega_j(x)$ is the sweep width of sensor j at x in R (see [3]). We assume that for some x, $U_1\omega_1(x) \neq U_2\omega_2(x)$.

To simplify the notation in the following, we assume that the local effectiveness functions b_j and sweep widths ω_j are explicitly independent of the location x in R. The modifications required to allow the dependence of these functions on position will be clear.

The class of search plans Ψ considered here is now described. A member of Ψ specifies two search effort density functions, each of which is a function of search time and location only. Thus, a plan in Ψ is a pair (m_1, m_2) where, for $j=1, 2$, m_j is a function defined on $R \times [0, \infty)$ such that $m_j(x, s)$ is the search effort density function for sensor j at x by search time s. We assume, for $j=1, 2$,

$$m_j(x, 0)=0$$

$$m_j(x,\cdot) \text{ is nondecreasing for } x\epsilon R.$$

We also assume that the search with sensors 1 and 2 takes place at constant speeds $U=U_1$ and $V=U_2$, respectively, so that

$$\frac{1}{U}\int_R m_1(x, s)dx+\frac{1}{V}\int_R m_2(x, s)dx=s. \tag{2}$$

The probability of having found the target by search time s is

$$P(m_1, m_2, s) \equiv \int_R f(x)[1-\bar{b}_1(m_1(x, s))\bar{b}_2(m_2(x, s))]dx, \tag{3}$$

where $\bar{b}_j=1-b_j$ for $j=1, 2$. The mean time to find the target using plan $(m_1, m_2)\ \epsilon\Psi$ is

$$\mu(m_1, m_2) \equiv \int_0^\infty sP(m_1, m_2, ds)=\int_0^\infty [1-P(m_1, m_2, s)]ds.$$

We minimize μ over plans in Ψ by maximizing P at each search time s over plans in Ψ. In particular, we find (m_1^*, m_2^*) in Ψ such that for any (m_1, m_2) in Ψ,

$$s' \geqslant s \text{ whenever } P(m_1, m_2, s') \geqslant P(m_1^*, m_2^*, s). \tag{4}$$

It then follows that for this (m_1^*, m_2^*),

$$\mu(m_1^*, m_2^*) \leqslant \mu(m_1, m_2). \tag{5}$$

3. OPTIMAL ALLOCATION OF EFFORT

In this section we construct an optimal plan (m_1^*, m_2^*) in Ψ. We appeal to a form of Theorem 3.2 in [8], stating it below, without proof.

Let R be a region in Euclidean n-space, and let G be the set of pairs of nonnegative functions (g_1, g_2) defined on R. For x in R, let $e(x, \cdot, \cdot)$ and $c(x, \cdot, \cdot)$ be real-valued functions defined on $[0, \infty) \times$

$[0, \infty)$, and for (g_1, g_2) in G let

$$E(g_1, g_2) = \int_R e(x, g_1(x), g_2(x))\, dx$$

$$C(g_1, g_2) = \int_R c(x, g_1(x), g_2(x))\, dx.$$

Define

$$\Delta(x) = \{q \mid q = c(x, z_1, z_2) \text{ for some } z_1, z_2 \geq 0\}.$$

THEOREM (from [8]): Suppose that for each $x \epsilon R$ and $q \epsilon \Delta(x)$ we may define $(\varphi_1(q), \varphi_2(q))$, such that

(6)
$$\text{(i)}\quad \varphi_1(q) \geq 0, \text{and} \varphi_2(q) \geq 0,$$

$$\text{(ii)}\quad c(x, \varphi_1(q), \varphi_2(q)) = q,$$

$$\text{(iii)}\quad e(x, \varphi_1(q), \varphi_2(q)) = \max\{e(x, z_1, z_2) \mid z_1, z_2 \geq 0 \text{ and } c(x, z_1, z_2) = q\} \equiv \xi(x, q).$$

Suppose for $x \epsilon R$, $\Delta(x)$ is an interval with lower and upper end-points $l(x)$ and $u(x)$ (possibly infinite), respectively and that $\xi'(x, \cdot)$, the derivative of $\xi(x, \cdot)$, exists and is Riemann integrable. Then if there exists a $k > 0$ and a function Q defined on R such that

(7)
$$1 \leq k\xi'(x, q) \qquad \text{for } l(x) < q \leq Q(x)$$

$$1 \geq k\xi'(x, q) \qquad \text{for } Q(x) < q < u(x),$$

it follows that $g^* = (g_1^*, g_2^*)$, where

$$g_j^*(x) \equiv \varphi_j(Q(x)), \qquad j = 1, 2,$$

has the property that for any $(g_1, g_2) \epsilon G$

$$E(g_1, g_2) \geq E(g_1^*, g_2^*) \text{ implies } C(g_1, g_2) \geq C(g_1^*, g_2^*).$$

To apply this theorem we now construct (m_1^*, m_2^*) in Ψ such that for each $s \geq 0$ $g_1^* = m_1^*(\cdot, s)$ and $g_2^* = m_2^*(\cdot, s)$ satisfy the above inequalities. Let $s \geq 0$ be fixed. Let

(8)
$$e(x, z, y) = f(x)[1 - \bar{b}_1(z)\bar{b}_2(y)]$$

$$c(x, z, y) = z/U + y/V \qquad \text{for } x \epsilon R,\ z \geq 0,\ y \geq 0.$$

Then clearly $\Delta(x) = [0, \infty)$ for each x in R. Note that $E(g_1^*, g_2^*) \equiv P(g_1^*, g_2^*, s)$ in (3) and that the constraint (2) can be written $C(g_1^*, g_2^*) = s$.

We first find the functions ξ, φ_1, and φ_2 which satisfy (6). Next, we show that $\xi'(x, \cdot)$ exists and is Riemann integrable, and finally we establish the existence of k and Q for which the inequalities (7) hold. The conclusion of the theorem then implies (4).

By using (8), (iii) and (ii) of (6) become, respectively,

$$\xi(x, q) = f(x)[1 - \bar{b}_1(\varphi_1(q))\,\bar{b}_2(\varphi_2(q))]$$

$$= \max_{0 \le z \le Uq} f(x)[b_1(z) + b_2(V(q - z/U)) - b_1(z)b_2(V(q - z/U))]$$

and

$$\varphi_1(q)/U + \varphi_2(q)/V = q. \tag{9}$$

Define F_q on $[0, Uq]$ for $q > 0$ by

$$F_q(z) = b_1(z) + b_2(V(q - z/U)) - b_1(z)b_2(V(q - z/U)).$$

To obtain ξ, we maximize F_q over $[0, Uq]$.

Differentiating F_q on $(0, Uq)$, we have

$$F_q'(z) = UG(q, z)\bar{b}_1(z)\bar{b}_2(V(q - z/U))$$

where

$$G(q, z) = UJ_1(z) - VJ_2(V(q - z/U)), \qquad 0 < z < Uq. \tag{10}$$

The condition $F_q'(z) = 0$ then becomes $G(q, z) = 0$, i.e.,

$$UJ_1(z) = VJ_2(V(q - z/U)). \tag{11}$$

By (iii) or (iv) of (1), J_1 and J_2 are nonincreasing on $(0, Uq)$ for $q > 0$ so that as z increases in $(0, Uq)$, the left-hand side of (11) is nonincreasing while the right-hand side of (11) is nondecreasing.

If, for all $q > 0$, $z > 0$, Equation (11) has no solution, then G always has the same sign. On the other hand, if $G(q, z) = 0$ for some q, z such that $q > 0$, $0 < z < Uq$, consider the values $UJ_1(0)$ and $VJ_2(0)$. If $UJ_1(0) \ge VJ_2(0)$, define

$$q^* \equiv \frac{1}{U} \min \{z | z \ge 0 \text{ and } UJ_1(z) \le VJ_2(0)\}.$$

Then q^* has the property that $G(q, z) > 0$ if $0 < q < q^*$, $0 < z < Uq$ (i.e., (11) has no solution for these ranges of q and z), and, for each $q > q^*$, Equation (11) has a unique solution $\psi(q)$, $0 < \psi(q) < Uq$. A similar argument can be made in case $UJ_1(0) < VJ_2(0)$.

We summarize the four possibilities for G as follows:

(i) $G(q, z) > 0$ for all $q > 0, z > 0$.

(ii) $G(q, z) < 0$ for all $q > 0, z > 0$.

(iii) $UJ_1(0) \ge VJ_2(0)$ and $G(q, z) = 0$ for some q, z such that $q > 0$, $0 < z < Uq$.

For this case define $q^* \equiv \frac{1}{U} \min \{z \mid z \geq 0 \text{ and } UJ_1(z) \leq VJ_2(0)\}$. Then q^*satisfies

(12) (a) $G(q, z) > 0$ if $0 < q < q^*$, $0 < z < Uq$,

(b) for each $q > q^*$, $G(q, z) = 0$ has a unique solution $\psi(q)$, $0 < \psi(q) < Uq$.

(iv) $VJ_2(0) > UJ_1(0)$ and $G(q, z) = 0$ for some q, z such that $q > 0$, $0 < z < Uq$.

For this case define $\hat{q} \equiv \frac{1}{V} \min \{y \mid y \geq 0 \text{ and } VJ_2(y) \leq UJ_1(0)\}$. Then $\hat{q}$ satisfies

(a) $G(q, z) < 0$ if $0 < q < \hat{q}$, $0 < z < Uq$,

(b) for each $q > \hat{q}$, $G(q, z) = 0$ has a unique solution $\theta(q)$, $0 < \theta(q) < Uq$.

We note that if $q^* = 0$ in (iii) of (12), then (a) of (iii) is vacuous and for $q > 0$, $G(q, z) = 0$, and hence $F_q'(z) = 0$, has a solution $\psi(q)$ which lies in $(0, Uq)$. Similarly, if $\hat{q} = 0$ in (iv) of (12), then (a) of (iv) is vacuous and for $\hat{q} > 0$, $G(q, z) = 0$ has a solution $\theta(q)$ which lies in $(0, Uq)$.

Since the sign of $F_q'(z)$ is determined by $G(q, z)$, in case (i) of (12), $F_q'(z) > 0$ for all $z > 0$ so that F_q is strictly increasing on $(0, Uq)$ for all $q > 0$. Thus, in this case, F_q has its maximum at $z = Uq$. Similarly, in case (ii) of (12), $F_q'(z) < 0$ for all $z > 0$ and F_q has its maximum at $z = 0$ for all $q > 0$.

In case (iii) of (12), we have for $0 < q \leq q^*$ that F_q is maximized at $z = Uq$ (as in case (i)), while for $q > q^*$, it follows that

$$G(q, z) > G(q, \psi(q)) = 0 \qquad \text{if } Uq^* < z < \psi(q)$$

and

$$G(q, z) < G(q, \psi(q)) = 0 \qquad \text{if } z > \psi(q).$$

Thus F_q is strictly increasing for $z < \psi(q)$ and strictly decreasing for $z > \psi(q)$, i.e., $\psi(q)$ maximizes F_q for $q > q^*$. A similar proof in case (iv) of (12) shows that for $0 < q \leq \hat{q}$, F_q is maximized at $z = Vq$, and for $q > \hat{q}$, $\theta(q)$ maximizes F_q.

Thus, for $q > 0$, the functions ξ, φ_1, and φ_2 of (6) are given as follows for each x in R:

Case (i): $\varphi_1(q) = Uq$, $\varphi_2(q) = 0$, $\xi(x, q) = e(x, Uq, 0)$.

Case (ii): $\varphi_1(q) = 0$, $\varphi_2(q) = Vq$, $\xi(x, q) = e(x, 0, Vq)$.

Case (iii): we have for $0 < q \leq q^*$,

(13) $$\varphi_1(q) = Uq, \ \varphi_2(q) = 0 \text{ and } \xi(x, q) = e(x, Uq, 0),$$

and for $q^* < q$,

$$\varphi_1(q) = \psi(q), \ \varphi_2(q) = V(q - \psi(q)/U) \text{ and } \xi(x, q) = e(x, \varphi_1(q), \varphi_2(q)).$$

Case (iv): we have for $0 < q \leq \hat{q}$,

$$\varphi_1(q) = 0, \ \varphi_2(q) = Vq \text{ and } \xi(x, q) = e(x, 0, Vq),$$

and for $\hat{q} < q$,

$$\varphi_1(q) = U(q - \theta(q)/V),\ \varphi_2(q) = \theta(q) \text{ and } \xi(x, q) = e(x, \varphi_1(q), \varphi_2(q)).$$

We now show that $\xi'(x, \cdot)$, the derivative of $\xi(x, \cdot)$, exists and is Riemann integrable. We consider case (iii) of (13) in detail and summarize all cases in (16). As a first step, we show that φ_1 is differentiable and nondecreasing. This is clear except for $q^* < q$. For this case $F'_q(\varphi_1(q)) = 0$ whence, from (11),

$$UJ_1(\varphi_1(q)) = VJ_2(V(q - \varphi_1(q)/U)), \text{ for } q > q^*, \tag{14}$$

and implicit differentiation yields

$$\varphi_1'(q) = \frac{UV^2 J_2'(V(q - \varphi_1(q)/U))}{U^2 J_1'(\varphi_1(q)) + V^2 J_2'(V(q - \varphi_1(q)/U))},\ q > q^*.$$

Thus φ_1' is well defined and $\varphi_1'(q) \geq 0$, $q > q^*$.

By the definition of ξ,

$$\xi'(x, \cdot) = f(x)[b_1'(\varphi_1(\cdot))\bar{b}_2(\varphi_2(\cdot))\varphi_1'(\cdot) + b_2'(\varphi_2(\cdot))\bar{b}_1(\varphi_1(\cdot))\varphi_2'(\cdot))].$$

Equation (14) yields $Ub_1'(\varphi_1(\cdot))\bar{b}_2(\varphi_2(\cdot)) = Vb_2'(\varphi_2(\cdot))\bar{b}_1(\varphi_1(\cdot))$ so that

$$\xi'(x, \cdot) = f(x)[b_1'(\varphi_1(\cdot))\bar{b}_2(\varphi_2(\cdot))(\varphi_1'(\cdot) + U\varphi_2'(\cdot)/V)].$$

Differentiating both sides of (9), we obtain $\varphi_1' + U\varphi_2'/V = U$ *and*

$$\xi'(x, \cdot) = Uf(x)\ b_1'(\varphi_1(\cdot))\bar{b}_2(\varphi_2(\cdot)). \tag{15}$$

Thus the existence of $\xi'(x, \cdot)$ is guaranteed.

Enumeration of all cases in (13) yields, for $q > 0$, $\xi'(x, q) = Uf(x)\beta(q)$, where β is defined as follows:

$$\beta(q) = \begin{cases} b_1'(Uq) & \text{if } q > 0 \text{ in case (i) or } 0 < q \leq q^* \text{ in case (iii) of (13)} \\ Vb_2'(Vq)/U & \text{if } q > 0 \text{ in case (ii) of (13) or } 0 < q \leq \hat{q} \text{ in case (iv) of (13)} \\ b_1'(\varphi_1(q))[1 - b_2(V(q - \varphi_1(q)/U))] & \text{if } q > q^* \text{ in case (iii) of (13) or if } q > \hat{q} \text{ in case (iv) of (13).} \end{cases} \tag{16}$$

Thus $\xi'(x, \cdot)$ exists, and since β is continuous, $\xi'(x, \cdot)$ is Riemann integrable. Note also that β is positive and strictly decreasing so that we may define β^{-1} on the range of β.

Finally, we show that there exist $k > 0$ and Q for which the inequalities (7) are satisfied. Once this is done, we will have satisfied the conditions of the theorem. Define, for each x in R and $k > 0$,

$$d(x, k) = \begin{cases} \beta^{-1}(1/kUf(x)), & \text{if } \beta(0) \geq \dfrac{1}{kUf(x)} > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Each choice of k corresponds to an amount of search time $A(k)$, where, for $k > 0$,

$$A(k) = \int_R d(x, k)dx.$$

As in [1], [2], and [4], we seek, for each $s > 0$, a k_s such that $A(k_s) = s$. For the case considered here, one may use the methods of [7] to show that by suitably restricting the domain of A, an inverse A^{-1} of A exists such that $k_s = A^{-1}(s)$ for each $s > 0$.

By the definition of d and the decreasing nature of β, the inequalities in (7) are satisfied for $k = k_s$ and $Q(x) = d(x, k_s)$. Define

$$\begin{aligned} &m^*(x, 0) = 0 \\ &m_j^*(x, s) = \varphi_j(d(x, k_s)), \qquad j = 1, 2, \text{ for } s > 0. \end{aligned} \tag{17}$$

Note that by (9), $C(m_1^*(\cdot, s), m_2^*(\cdot, s)) = s$ and that $(m_1^*, m_2^*) \in \Psi$. Then by the theorem, we have for any $(g_1, g_2) \in G$,

$$E(g_1, g_2) \geq E(m_1^*(\cdot, s), m_2^*(\cdot, s)) \text{ implies } C(g_1, g_2) \geq C(m_1^*(\cdot, s), m_2^*(\cdot, s)) = s. \tag{18}$$

Since $P(m_1^*, m_2^*, s) = E(m_1^*(\cdot, s), m_2^*(\cdot, s))$, (18) yields (4). From this (5) follows. By performing the above construction for each $s \geq 0$, we obtain an optimal search plan (m_1^*, m_2^*) in Ψ.

4. EXAMPLE

In this section we present an example of search for a stationary target in R, the real plane. Let the prior target location distribution be circular normal. We assume that one of the two search sensors has a deterministic sweep width and that the sweep width of the other sensor is stochastic. Sensors with stochastic sweep widths have been considered in [5], [6], and [8].

Let sensor 2 have deterministic sweep width $\omega > 0$ and search speed V. The *sweep rate* of sensor 2 is the product ωV of sweep width and search speed. Let the local effectiveness function be $b_2(z) = 1 - \exp(-\omega z)$, $z \geq 0$.

Let W be a random variable giving the value of sweep width for sensor 1. We assume that the value of sweep width, although unknown, is fixed throughout the search. For $z, w \geq 0$, we define $B_1(z, w)$ to be the probability of detecting the target at any point $x \in R$ with sensor 1 given that the target is located at x, that $W = w$, and that effort density z has been expended at x. We suppose that $B_1(z, w) = 1 - \exp(-wz)$ and that W is gamma distributed with scale parameter η and convolution parameter ν. Letting F denote the distribution function of W, one may check that

$$b_1(z) = \int_0^\infty B_1(w, z)F'(w)dw = 1 - [\eta/(\eta + z)]^\nu \qquad \text{for } z \geq 0.$$

We refer to the function B_1 as the *conditional local effectiveness function* for sensor 1. If U is the search speed for sensor 1, we define the product $U\nu/\eta$ of search speed and mean sweep width as the prior *mean sweep rate* for sensor 1. The choice of search sensor will be governed by comparison of sweep rate with mean sweep rate.

The prior target location distribution density function is given, in polar coordinates, by

$$f(\rho, \theta) = \frac{1}{2\pi\sigma^2} e^{-\rho^2/2\sigma^2}, \rho > 0, 0 \leq \theta < 2\pi,$$

where σ is a positive constant. Since the form of f is independent of θ, we write $f(\rho)$ for $f(\rho, \theta)$.

As a specific example for later use, Table 1 presents hypothetical parameter values for the search sensors and target location distribution. It is assumed that search effort is measured in track length (miles) so that the search effort density has units of reciprocal miles.

TABLE 1 *Example of Parameter Values*

Target location: standard deviation, $\sigma = 2$ miles
Sensor 1: scale parameter, $\eta = 1$ (miles)$^{-1}$ convolution parameter, $\nu = 2$ search speed, $U = 1$ knot
Sensor 2: sweep width, $\omega = 0.5$ miles search speed, $V = 2$ knots

One can check that b_1 and b_2 above satisfy (i) through (iii) of (1), and we have

$$e(\rho, z, y) = f(\rho)\left[1 - e^{-\omega y}\left(\frac{\eta}{\eta + z}\right)^{\nu}\right]$$

$$c(\rho, z, y) = z/U + y/V \qquad \text{for } \rho > 0,\ z, y \geq 0.$$

Then ξ is given as follows:

(i) If $U\nu < \eta V\omega$, then $\varphi_1(q) = 0$, $\varphi_2(q) = Vq$, and $\xi(\rho, q) = f(\rho)[1 - e^{-\omega V q}]$, $q > 0$.

(ii) If $U\nu > \eta V\omega$, we have

for
$$0 < q \leq \frac{1}{U}\left(\frac{U\nu}{V\omega} - \eta\right),$$

$$\varphi_1(q) = Uq,\ \varphi_2(q) = 0, \text{ and } \xi(\rho, q) = f(\rho)\left[1 - \left(\frac{\eta}{\eta + Uq}\right)^{\nu}\right];$$

for
$$q > \frac{1}{U}\left(\frac{U\nu}{V\omega} - \eta\right),$$

$$\varphi_1(q) = \frac{U\nu}{V\omega} - \eta,\ \varphi_2(q) = V\left[q - \frac{1}{U}\left(\frac{U\nu}{V\omega} - \eta\right)\right],\ \xi(\rho, q) = [1 - Ke^{-\omega V q}],$$

where

$$K = \left(\frac{\eta V\omega}{U\nu}\right)^{\nu} \exp\left\{\nu - \frac{\eta V\omega}{U}\right\}.$$

In case (i) above, only the deterministic sensor is used. Thus $\beta(q) = \omega e^{-\omega V q}$ in formula (16) and

$$\beta^{-1}(q) = \frac{1}{\omega V} \ln\left(\frac{\omega}{q}\right),\ \omega \geq q > 0.$$

One may check that in this case the optimal allocation reduces to the one given by Koopman in [5]. (If the value for V in Table 1 is replaced by $V = 5$ knots, then a specific example for this case is obtained.)

For case (ii) we have

$$\xi'(\rho, q) = \begin{cases} Uf(\rho)\dfrac{\nu\eta^{\nu}}{(\eta + Uq)^{\nu+1}}, & 0 < q < \dfrac{1}{U}\left(\dfrac{U\nu}{V\omega} - \eta\right) \\ \omega VKf(\rho)e^{-\omega Vq}, & q \geq \dfrac{1}{U}\left(\dfrac{U\nu}{V\omega} - \eta\right) \end{cases}$$

and

$$d(\rho, k) = \begin{cases} \dfrac{1}{U}\left[\left(\dfrac{\nu\eta^{\nu}}{r(\rho, k)}\right)^{1/(\nu+1)} - \eta\right] & \text{if } \dfrac{\nu}{\eta} \geq r(\rho, k) > L \\ \dfrac{1}{\omega V}\ln\left(\dfrac{\omega VK}{U\, r(\rho, k)}\right) & \text{if } L \geq r(\rho, k) > 0 \\ 0 & \text{otherwise,} \end{cases} \tag{19}$$

where $L = \nu\eta^{\nu}(V\omega/U\nu)^{\nu+1}$ and r is given by

$$r(\rho, k) = \frac{2\pi\sigma^2}{kU}\, e^{\rho^2/2\sigma^2},\ k > 0. \tag{20}$$

The first inequality in (19) becomes, upon substituting for $r(\rho, k)$

$$2\sigma^2 \ln\left(\frac{kUL}{2\pi\sigma^2}\right) < \rho^2 \leq 2\sigma^2 \ln\left(\frac{\nu kU}{2\pi\sigma^2\eta}\right) \tag{21}$$

while the second becomes

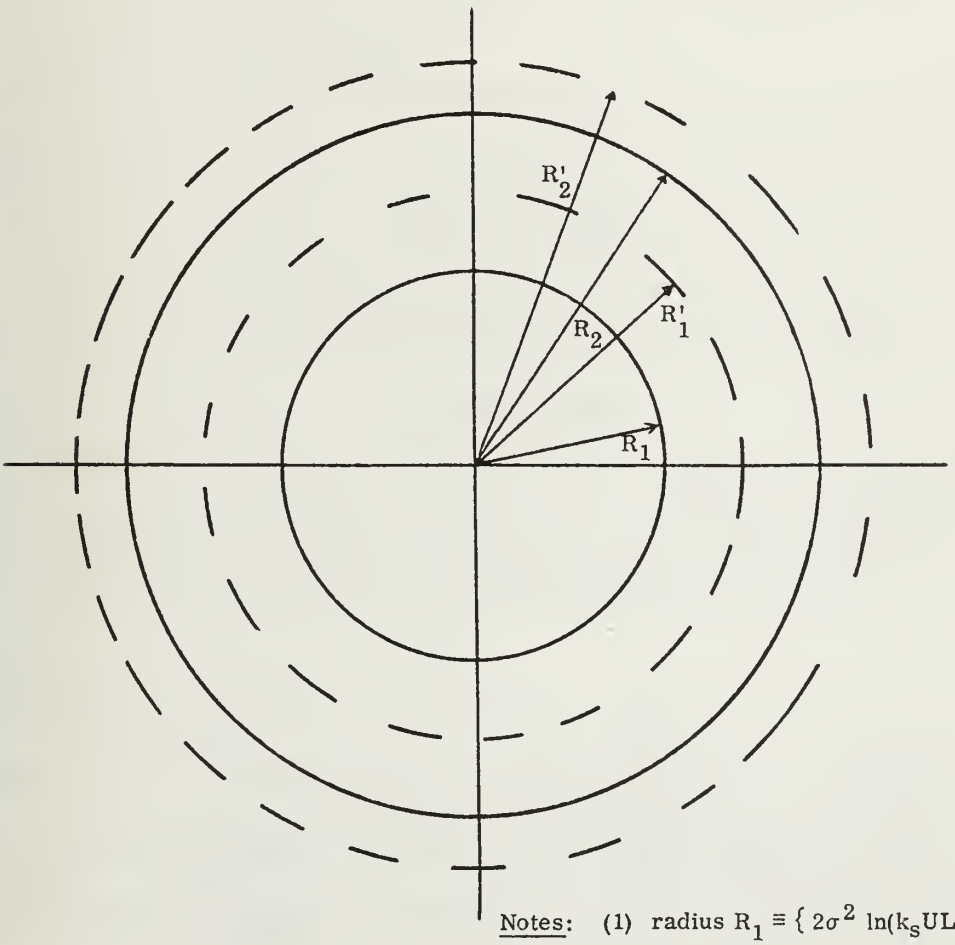
$$0 < \rho^2 < 2\sigma^2 \ln\left(\frac{kUL}{2\pi\sigma^2}\right). \tag{22}$$

One may now evaluate A on $[2\pi\sigma^2\eta/U\nu, \infty)$; numerical methods must be used to find A^{-1}. Define $k_s = A^{-1}(s)$ for $s > 0$ and using (17) and (19) we obtain

$$m_1^*(\rho, s) = \begin{cases} \left(\dfrac{\nu\eta^{\nu}}{r(\rho, k_s)}\right)^{1/(\nu+1)} - \eta & \text{if } 2\sigma^2 \ln\left(\dfrac{k_sUL}{2\pi\sigma^2}\right) < \rho^2 \leq 2\sigma^2 \ln\left(\dfrac{\nu k_sU}{2\pi\sigma^2\eta}\right) \\ \dfrac{U\nu}{V\omega} - \eta & \text{if } 0 < \rho^2 \leq 2\sigma^2 \ln\left(\dfrac{k_sUL}{2\pi\sigma^2}\right) \\ 0 & \text{if } \rho^2 > 2\sigma^2 \ln\left(\dfrac{\nu k_sU}{2\pi\sigma^2\eta}\right) \end{cases}$$

(23)
$$m_2^*(\rho, s) = \begin{cases} 0 & \text{if } 2\sigma^2 \ln\left(\frac{k_s UL}{2\pi\sigma^2}\right) < \rho^2 \leq 2\sigma^2 \ln\left(\frac{\nu k_s U}{2\pi\sigma^2\eta}\right) \\ \frac{1}{\omega} \ln\left(\frac{\omega VK}{Ur(\rho, k_s)}\right) - \frac{\nu}{\omega} - \frac{\eta V}{U} & \text{if } 0 < \rho^2 \leq 2\sigma^2 \ln\left(\frac{k_s UL}{2\pi\sigma^2}\right) \\ 0 & \text{if } \rho^2 > 2\sigma^2 \ln\left(\frac{\nu k_s U}{2\pi\sigma^2\eta}\right). \end{cases}$$

Examination of the formulas for m_1^* and m_2^* shows that the search is conducted in two regions, a circular disk centered at the origin and an annular region bounding the disk as shown in Figure 1. Search in the disk with radius R_1 is conducted using sensor 2 while search in the annular region bounded by the circles R_1 and R_2 is conducted using sensor 1. As the search progresses, the radius R_1 increases while the difference $R_2 - R_1$ decreases. Thus the progressive search regions consist of an expanding disk with a narrowing annulus attached.



Notes: (1) radius $R_1 \equiv \{ 2\sigma^2 \ln(k_s UL/2\pi\sigma^2)\}^{\frac{1}{2}}$.

(2) radius $R_2 \equiv \{ 2\sigma^2 \ln(\nu k_s U/2\pi\sigma^2\eta)\}^{\frac{1}{2}}$.

FIGURE 1. Progressive search regions

If we fix ρ and let the search time s increase, then from (23) we see that effort is exerted by sensor 1 at a point (ρ, θ) until $Uv/V\omega - \eta$ effort has been applied. Succeeding effort is exerted at the point only by sensor 2. Hence, if the search has progressed optimally to time s, and more effort becomes available, the search expands as shown by the dashed curves in Figure 1. The circles R_1 and R_2 are expanded to R_1' and R_2', respectively, and effort is applied as follows. Search in the disk with radius R_1' is conducted by using only sensor 2 and in the annulus bounded by the circles R_1' and R_2' by using only sensor 1.

For the specific parameter values given in Table 1, the progressive search region boundaries R_1 and R_2 are shown in Table 2 for different values of the broad search time s. The values for $k_s = A^{-1}(s)$ were obtained by numerically inverting the function A. Note that in this case, the radius R_1 remains negative until broad search time $s \approx 23$ hours so that no search is performed using sensor 2 for nearly 2 days. Thereafter, the search is conducted in the manner described above.

TABLE 2 *Search Regions for Specific Example*

(note: Parameter values are shown in Table 1)

Search time s (hr)	Value $A^{-1}(s) = k_s$ (hr)	Radius R_1 (mi)	Radius R_2 (mi)
1	20.3	—	1.95
6	38.8	—	3.00
12	59.1	—	3.52
24	104.0	0.521	4.11
36	153.5	1.84	4.47
48	207.8	2.41	4.74

REFERENCES

[1] DeGuenin, Jacques, "Optimum Distribution of Effort: An Extension of the Koopman Basic Theory," Operations Research, *9*, 1–7 (1961).

[2] Karlin, S., Mathematical Methods and Theory in Games, Programming, and Economics, Vol. II, The Theory of Infinite Games (Addison-Wesley Publishing Company, Inc., Reading, 1959).

[3] Koopman, B. O., "The Theory of Search: Part II, Target Detection," Operations Research, *4*, 503–531 (1956).

[4] Koopman, B. O., "The Theory of Search: Part III, The Optimum Distribution of Searching Effort," Operations Research, *5*, 613–626 (1957).

[5] Koopman, B. O., Search and Screening, Operations Evaluation Group Report 56, Office of the Chief of Naval Operations, Navy Department, Washington, D.C. (1946).

[6] Richardson, Henry R. and Barry Belkin, "Optimal Search With Uncertain Sweep Width," Operations Research, *20*, 764–784 (1972).

[7] Stone, Lawrence D. and Jerome A. Stanshine, "Optimal Search Using Uninterrupted Contact Investigation," SIAM J. Appl Math., *20*, 241–263 (1971).

[8] Stone, Lawrence D., Jerome A. Stanshine, and Carl A. Persinger, "Optimal Search in the Presence of Poisson-Distributed False Targets," SIAM J. Appl. Math, *23*, (1972).

BOUNDS ON THE AVAILABILITY FUNCTION

Richard W. Butterworth

and

Tor Nikolaisen*

Naval Postgraduate School
Monterey, California

ABSTRACT

This paper gives bounds on the availability function for an alternating renewal process with exponential failure and general repair times. A bound on the error is also given. Several of the bounds with greatest practical consequence are worked out and illustrated. Repair distributions for which a lower bound on availability is easily computed are gamma (integer shape parameter), log normal, and Weibull. Finally, some simulation results for log normal repair versus gamma repair are given.

INTRODUCTION AND SUMMARY

The standard model in the reliability literature for a one unit repairable system is the alternating renewal process. This model is appropriate for a system (or subsystem) which begins in an operative condition and operates for some random time, begins repair whenever failure occurs and when the repair is completed, is fully restored to an operating condition where it remains until failure followed by repair, . . . etc. The availability of such a system is defined as the probability that the system is operative.

The basic problem treated here is to obtain bounds on the availability function for the case of exponential failure and general repair distributions. Both lower and upper bounds are given and a conservative estimate of their accuracy is provided. The bounds are worked out for the gamma, Weibull, and log normal repair distributions.

More complex models for availability are numerous in the literature; however, they usually obtain steady-state results or transforms of the availability function at best. The cases which can be solved exactly usually require every random time to have the exponential distribution. See Shooman [5] for example. While the exponential failure law is widely accepted, the assumption of an exponential repair law is not. As noted in Arinc [1], the log normal is frequently believed to describe repair times. The model given here then deals with some more realistic repair time distributions, necessarily in a relatively simple situation.

An application of the present work is to give a correct account of the availability as a function of time in lieu of the steady-state value which must frequently be taken as a constant value in practical situations. It may, in fact, justify the use of steady-state values. Another application which is briefly discussed is that of replacing a very regular repair time distribution by the constant repair time distribution for which the availability can be found exactly. Finally, some simulation results for log normal repair are given. These suggest that the gamma distribution can be used in lieu of log normal for computing the availability function.

---

* Presently at Naval Command Vestlandet, Mathopen, Norway.

Some numerical examples are given to illustrate the practicality of the bounds presented. While a computer was used for some calculations, the programming effort required was minimal.

BASIC RESULTS

An alternating renewal process is specified by two sequences, $\{X_1, X_2, \ldots\}$ and $\{Y_1, Y_2, \ldots\}$, of random variables; the first sequence represents the durations of operation between failures and the second the durations of repair between failures. Thus, the process is in an operative state during $[0, X_1)$, under repair during $[X_1, X_1+Y_1)$, operative during $[X_1+Y_1, X_1+Y_1+X_2), \ldots$ etc. (Note that we could begin in a repair state if desired.) The operative times $X_1, X_2, \ldots$, will be given the exponential distribution, parameter λ, and the repair times $Y_1, Y_2, \ldots$, will have an unspecified distribution G, with some finite positive mean. All random times are assumed to be jointly independent. Thus we have,

$$P(X_n \leq t) = 1 - e^{-\lambda t} \qquad n = 1, 2, \ldots$$

$$P(Y_n \leq t) = G(t) \qquad n = 1, 2, \ldots$$

(1)

$$X_1, Y_1, X_2, Y_2, \ldots \text{ independent random variables.}$$

Within the assumptions of (1), we wish to compute the availability function $A(t)$, the probability of being in an operative state at time t. A precise definition would be

$$A(t) = P\left(\sum_{k=1}^{n} (X_k+Y_k) \leq t < \sum_{k=1}^{n} (X_k+Y_k) + X_{n+1} \qquad \text{for some } n \geq 0\right),$$

where empty summations are taken as zero in the probability statement.

From our convention of beginning operative at time = 0, we have that $A(0) = 1$. Due to the assumptions in (1), we know that $A(\infty) = \lim_{t\to\infty} A(t) = \dfrac{\nu}{\nu+\lambda}$, where ν^{-1} = mean repair time under the distribution G (as shown in [2]). For obvious reasons, we will call λ the failure rate and ν the repair rate (even when G is not an exponential). Note that the steady-state value of availability depends only on the first moment of G.

Using script letters to denote Laplace-Stieltjes transforms, so $\mathscr{A}(s) = \int_0^\infty e^{-sx}\, dA(x)$, it is easily established that (see [2])

$$\mathscr{A}(s) = \frac{1 - \dfrac{\lambda}{\lambda+s}}{1 - \dfrac{\lambda}{\lambda+s}\mathscr{G}(s)},$$

which simplifies to

(2)
$$\mathscr{A}(s) = \frac{1}{1+r\mathscr{G}_e(s)} = \sum_{n=0}^{\infty} (-r\mathscr{G}_e(s))^n,$$

where

$$r = \lambda/\nu$$

$$G_e(t) = \nu \int_0^t (1 - G(x))\, dx$$

and

$$\mathcal{G}_e(s) = \frac{\nu}{s}(1 - \mathcal{G}(s)).$$

Assuming $r < 1$, the series in (2) is valid and gives, on inversion,

$$A(t) = 1 - rG_e(t) + r^2 G_e^* G_e(t) - r^3 G_e^* G_e^* G_e(t) + \ldots, \tag{3}$$

where * denotes a convolution, and we assume $r = \lambda/\nu < 1$.

The distribution G_e appearing in (3) is well known in renewal theory as the equilibrium excess distribution for a renewal process with inter-event distribution G. See [3] or [2] for a renewal-theoretical interpretation. Unfortunately, no such interpretation applies to the present work.

The expansion given by (3) can be terminated to give either upper or lower bounds on $A(t)$; an estimate of the error shows that arbitrarily high accuracy can be obtained by taking sufficiently many terms. First notice that the terms in (3) are, in absolute value, monotonic decreasing, for any t. Now define the following functions:

$$\text{for } n \text{ odd}, L_n(t) = \sum_{k=0}^{n} r^k G_e^{k*}(t);$$

$$\text{for } n \text{ even}, U_n(t) = \sum_{k=0}^{n} r^k G_e^{k*}(t).$$

Using the above equality and (3), we have that

$$\text{for } n \text{ odd: } L_n(t) \leqq A(t), \text{ and } A(t) - L_n(t) \leqq r^{n+1};$$

$$\text{for } n \text{ even: } U_n(t) \geqq A(t), \text{ and } U_n(t) - A(t) \leqq r^{n+1} \tag{4}$$

$$L(t) = L_1(t), U(t) = U_2(t).$$

The upper and lower bounds U_n and L_n can be made to have an arbitrarily small error by choosing n sufficiently large, since $r < 1$. The bounds of greatest practical interest are of course L and U. The restriction $r < 1$ is the same as $A(\infty) > 0.5$, since $A(\infty) = 1/(1+r)$. For the desirable case of $A(\infty) = 90$ percent, we have $r = 1/9$ and for $A(\infty) = 98$ percent, $r = 1/49$. Then, for example, the lower bound $L(t) = L_1(t) = 1 - \lambda \int_0^t (1 - G(x))dx$ has a maximum error of $r^2 = 0.0123$ or $r^2 = 0.000416$ for $A(\infty) = 90$ or 98 percent, respectively.

The first (nontrivial) upper bound is $U(t) = U_2(t)$. As an example, this bound is easily computed for a repair distribution G, which is a gamma with an integer shape parameter. Numerical examples are given in the following section. The maximum error of U for any repair distribution G is r^3 which is 0.00138 or 0.0000085 for $A(\infty) = 90$ or 98 percent, respectively.

The bounds U_n and L_n are usually cumbersome to compute for $n \geqslant 3$ unless the distribution G_e takes on a simple form. One can, however, define a lower bound whose error is at most r^3 involving

no more computation than that involved in U, namely $U(t)-r^3=1-rG_e(t)+r^2G_e^{2*}(t)-r^3$, for which we have

$$0 \leqq A(t)-(U(t)-r^3) \leqq r^3. \tag{5}$$

By combining the lower bound L and (5), we define a revised version of L, namely $L'(t)=\text{Max}\,(L(t), U(t)-r^3)$, for which we have

$$0 \leqq A(t)-L'(t) \leqq r^3. \tag{6}$$

SPECIFIC REPAIR DISTRIBUTIONS

The bounds L and in a few cases U and L' can be easily computed for many plausible repair distributions. For a discussion of some of these cases and a thorough treatment of the exact computation of availability for gamma distributed repair, see Nikolaisen [4].

One case for which all three bounds L, U, and L' are easily computed is when G is a gamma distribution with integer shape parameter. If α (integer) is the shape parameter and $\beta>0$ the scale parameter, we have

$$G(t)=1-\sum_{j=1}^{\alpha}\frac{(\beta t)^{j-1}}{(j-1)!}e^{-\beta t}$$

so

$$G_e(t)=1-\sum_{j=1}^{\alpha}\frac{(\beta t)^{j-1}}{(j-1)!}e^{-\beta t}\frac{(\alpha-j+1)}{\alpha}.$$

Recall that we are always taking the failure distribution to be exponential with parameter λ. Since $r=\frac{\lambda}{\nu}=\frac{\lambda\cdot\alpha}{\beta}$, the lower bound L becomes

$$L(t)=1-\frac{\lambda\alpha}{\beta}+\frac{\lambda}{\beta}\sum_{j=1}^{\alpha}\frac{(\beta t)^{j-1}}{(j-1)!}e^{-\beta t}(\alpha-j+1).$$

The upper bound U is also a straight-forward computation with

$$U(t)=1-\frac{\lambda\alpha}{\beta}\left(1-\frac{\lambda\alpha}{\beta}\right)G_e(t)-\left(\frac{\lambda}{\beta}\right)^2\sum_{k=1}^{\alpha}\sum_{j=1}^{\alpha}(\beta t)^{j+k+1}\frac{(\alpha-j+1)}{(j+k-1)!}e^{-\beta t}.$$

The revised version $L'(t)$ based on $U(t)$ is simply

$$\text{Max}\left(L(t), U(t)-\left(\frac{\lambda\alpha}{\beta}\right)\right).$$

Figure 1 shows the bounds, L, L', U, and the exact availability A for $\alpha=4$, $\beta=18.0$ and $\lambda=0.5$. The exact availability was found by inverting the transform equation (2). See [4] for a closed form expres-

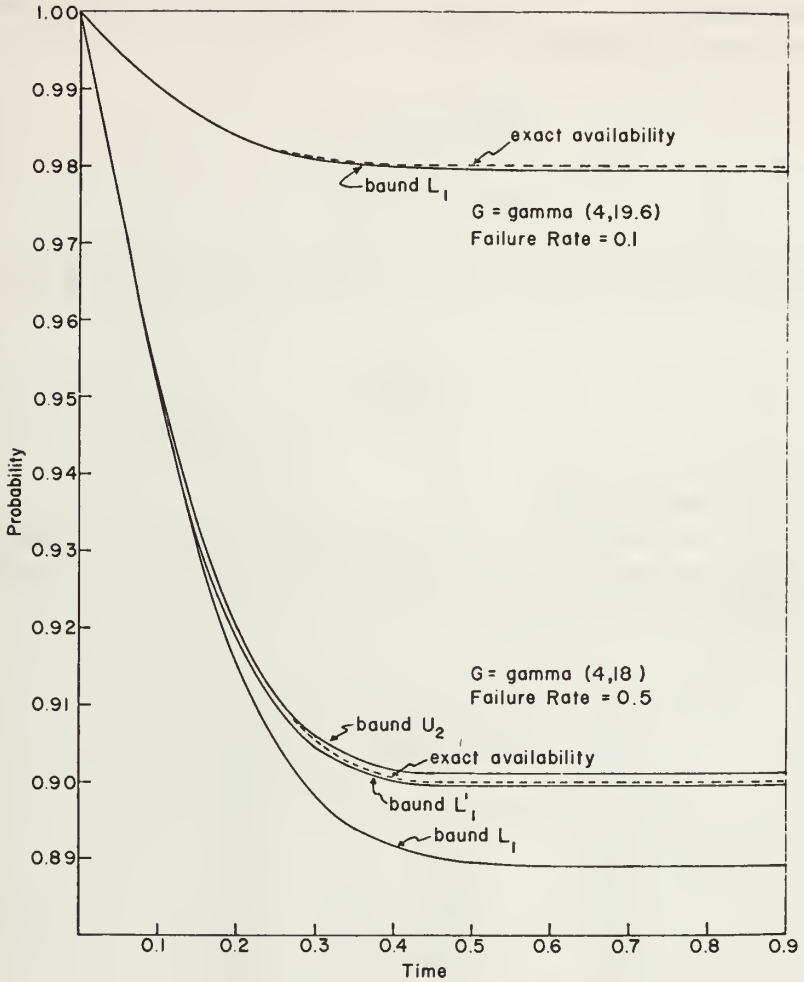


FIGURE 1. Illustration of bounds on availability

sion. The repair rate ν equals 4.5 so the steady-state value $A(\infty) = 90$ percent. This graph is typical for α less than, say 10, with β such that the repair rate ν remains fixed at 4.5.

Figure 1 also shows the lower bound L and the exact availability for $\alpha = 4$, $\beta = 19.6$ and $\lambda = 0.1$ so $\nu = 4.9$ and the steady-state availability $A(\infty) = 98$ percent. The accuracy of the bounds is significantly improved for these values; in fact, both L' and U are indistinguishable from the exact availability on this graph.

Two more specific repair distributions for which the lower bound L can be computed are the log normal and the Weibull. For the log normal with parameters μ and σ^2 (not the mean and variance) we have

$$G(t) = \Phi((\ln(t) - \mu)/\sigma)$$

and

$$G_e(t) = \Phi((\ln(t) - \mu - \sigma^2)/\sigma) + e^{-\mu-\sigma^2/2}t\Phi((\ln(t) - \mu)/\sigma),$$

where Φ is the standard normal distribution function. The value of r becomes, in this case,

$$r = \lambda e^{\mu+\sigma^2/2}.$$

The Weibull distribution can also be integrated to give the lower bound L, in terms of the gamma distribution function. Namely, we have that

$$G(t) = 1 - e^{-\beta t^{\alpha}}$$

and

$$G_e(t) = \Gamma(t^{\alpha}),$$

where Γ is the gamma distribution function with shape parameter $1/\alpha$ and scale parameter β. In this case, we have

$$r = \lambda \frac{\gamma(1/\alpha + 1)}{\beta^{1/\alpha}},$$

where γ denotes the usual gamma function.

Table 1 given shows the lower bound L for the log normal and Weibull distributions. The choice of parameters for these examples is based on having a steady-state value of availability, $A(\infty)$, of 90 percent and a coefficient of variation less than 1.0.

TABLE 1

Time (t)	Log normal repair[a] lower bound $L(t)$	Weibull repair[a] lower bound $L(t)$
0.00	1.000	1.0000
0.05	0.9752	0.9765
0.10	0.9533	0.9567
0.15	0.9363	0.9407
0.20	0.9239	0.9280
0.25	0.9150	0.9181
0.30	0.9086	0.9105
0.40	0.9005	0.9005
0.50	0.8960	0.8949
0.75	0.8913	0.8899
1.00	0.8898	0.8891

[a]Both cases have $\lambda = 0.5$, $\nu = 4.5$, $A(\infty) = 90$ percent and repair coefficient of variation = 0.805

APPLICATIONS

One obvious application of the above bounds is to account for the availability as a function of time in lieu of the steady-state value which is frequently used. The bounds may even justify the use of steady-state values in particular applications.

As another application, suppose the repair process being studied is very regular. As the availability is easily computed exactly for exponential failure and constant repair, the question that arises is how regular do repair times need to become for the assumption of constant repair to approximate availability adequately. This question is studied by using the gamma distribution with a large shape parameter as a model for regular repair times. One can then compare the bounds on availability in this case with the exact solution for constant repair. Figure 2 illustrates this comparison. The shape param-

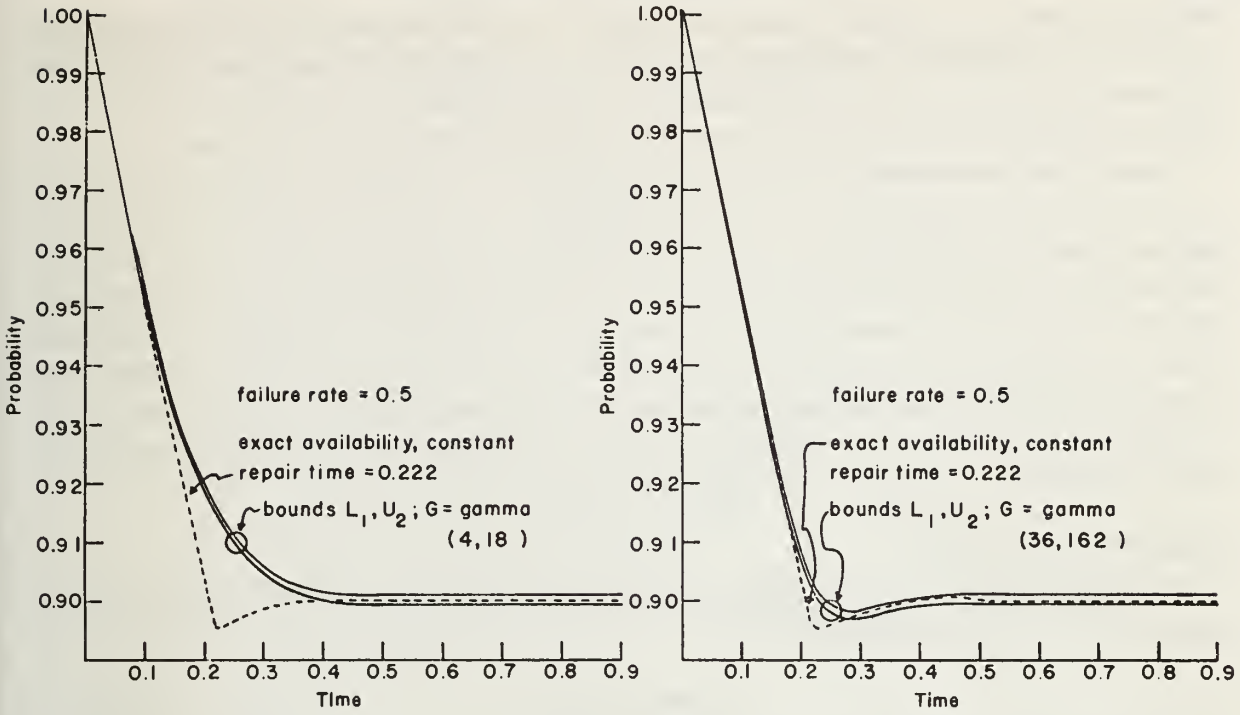


FIGURE 2. Availability for gamma repair vs constant repair

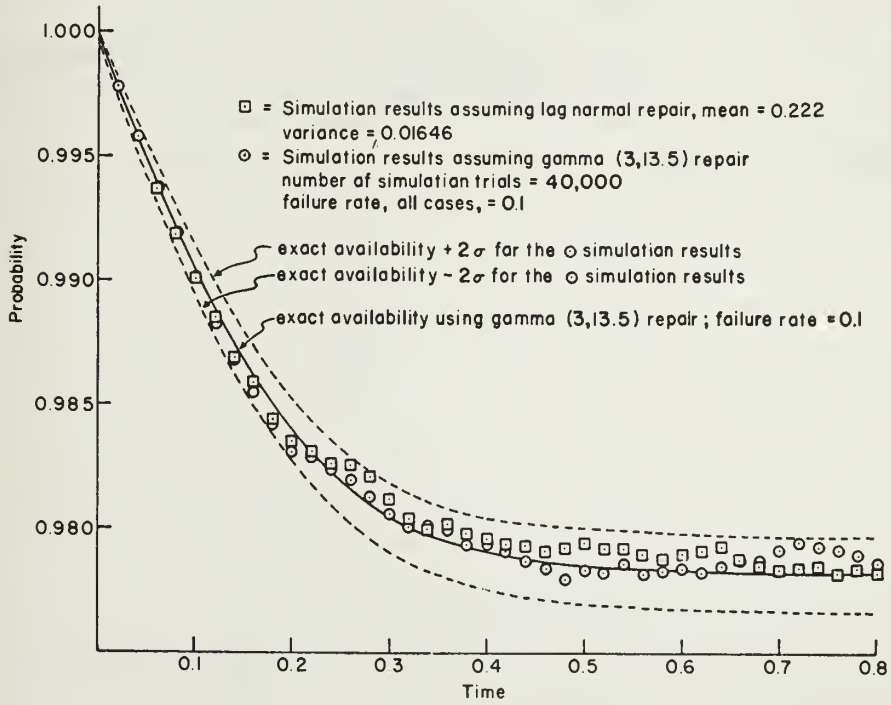


FIGURE 3. Comparison of gamma to log normal repair assumptions

eter is 4 and 36 in the two plots shown, respectively. The scale parameter was chosen to make the repair rate 4.5 in both cases. A useful measure of regularity is the coefficient of variation. This coefficient is 1/2 and 1/6, respectively, in the plots shown. Computing experience not shown here verifies that the same general situation prevails when the steady-state availability is 98 instead of 90 percent.

Figure 3 shows the results of 40,000 simulations by using first the log normal repair distribution and then the gamma repair distribution with shape parameter 3. The other parameters were chosen so that both repair distributions had the same mean and variance; the repair rate $\nu = 4.5$ and the repair time variance = 0.01646. The failure rate was = 0.1 so that the steady state $A(\infty) = 97.83$ percent. Also shown is the exact value of availability for the gamma repair distribution used. The $\pm 2\sigma$ lines refer to the gamma repair simulation averages. That is, these simulation outcomes are normally distributed with the exact availability as mean and standard deviation σ as indicated. The log normal simulation results are seen to lie in this band as well. Within the limits of stochastic modelling, the use of a gamma in place of log normal repair distribution may well be acceptable.

ACKNOWLEDGMENTS

The referees provided helpful suggestions in the presentation of these results. This research was supported by the Office of Naval Research.

REFERENCES

[1] Arinc Research Corporation, *Reliability Engineering* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1964), pp. 400–405.

[2] Barlow, R. and Frank Proschan, *Mathematical Theory of Reliability* (John Wiley and Sons, New York, 1965), pp. 74–82.

[3] Cox, D. R., *Renewal Theory* (Methuen & Co., Ltd., London, 1962), pp. 80–86.

[4] Nikolaisen, Tor Mikal, "Alternating Renewal Process with Exponentially Distributed Time to Failure, Bounds with Precision," Master's Thesis, Naval Postgraduate School, Monterey, California (Sept. 1971).

[5] Shooman, Martin L., *Probabilistic Reliability: an Engineering Approach* (McGraw-Hill, Inc., New York, 1968), pp. 336–360.

THE SINGLE SERVER QUEUE IN DISCRETE TIME– NUMERICAL ANALYSIS I

Marcel F. Neuts*

Purdue University

ABSTRACT

This is the first of a sequence of papers dealing with the computational aspects of the transient behavior of queues in discrete time.

It is shown that for a substantial class of queues of practical interest, a wealth of numerical information may be obtained by relatively unsophisticated methods.

This approach should prove useful in the analysis of unstable queues which operate over a limited time interval, but is by no means limited to such queues.

Mathematically the service unit is modeled in terms of a multivariate Markov chain, whose particular structure is used in iterative computation. Many important queue features may then be derived from the n-step transition probabilities of this chain.

1. INTRODUCTION

The theory of queues, and more generally that of stochastic models, suffer from the insufficient development of the interface between structural-analytic results on one hand and directly applicable numerical methods on the other hand.

The practical queue analyst tends to use of the extensive theoretical work on service systems only those rare steady-state results which are analytically simple. This is often done with little regard for the mathematical assumptions underlying these results. Moreover such results commonly do not answer the real questions one is facing in the design of a service facility, and in rare cases the measures of performance based on steady-state assumptions may actually be misleading. An example of a stable queue with rare arrivals of large groups of customers in which this is the case, is discussed in the third paper in this sequence.

While simulation techniques are widely used, their implementation requires a thorough understanding of the probability structure of the queue as well as very substantial computing funds. In many instances of simulation studies familiar to this author, the structure of the queue was incorrectly or insufficiently used and exorbitant computing times were reported. One recognizes that simulation is often the only resort in studying a complex system. However, for those models whose structure is mathematically well understood, it is desirable that algorithms making use of the existing theory be developed. It is obviously intellectually pleasant to be able to use one's understanding of the mathematical structure to obtain numerical results by efficient algorithms. There are however also many practical reasons for investigating exact, rather than Monte Carlo algorithms, wherever possible. Since this is not the place to discuss these at length, we mention the study of relatively rare events as an example. To obtain a good estimate of the probability that a very stable queue exceeds a certain bound may require long simulation runs, because the simulated paths will only rarely exhibit the event

* This research was supported by the National Science Foundation, grant GP 28650.

of interest. Queues which are nearly critical are also difficult to simulate, because of the substantial stochastic variability of the queue length and the related quantities. For these again, exact algorithms prove to be a useful alternative method of analysis.

The time involved in the preparation and testing of an exact algorithm is probably much greater than in a corresponding simulation study. This aspect must be taken into account in comparisons of the cost and the effectiveness of both approaches. For this reason only models of wide potential applicability should initially be studied by means of exact algorithms. There is furthermore a return which has particular value to the theoretical queue analyst. An algorithm is similar to a theorem in that its applicability is usually far greater than the original problem from which it sprung. As is the case for theorems, algorithms are also interesting because of their limitations. When the latter are recognized, they usually stimulate many questions related to approximations, algorithmic efficiency or structural theorems.

The purpose of this paper and the subsequent ones is to investigate a useful single server queue in detail. Before we consider the model specifically, we first discuss some difficulties and desiderata related to the numerical analysis of a much wider class of queueing problems.

a. The Finiteness of the Waiting Line

In reality, unbounded queues do not exist. The unbounded queue is strictly a mathematically convenient abstraction. By removing one "boundary" of the queue length process one obtains simpler stochastic processes. It further becomes possible to give an elegant treatment of the intuitive quality of stability of the queue.

In practical situations, there is either a finite waiting room, usually with loss of those customers who find the waiting space full, or else the buildup of the queue beyond a certain limit creates utter chaos and the object of the study is precisely the design of a sufficiently fast service unit or a sufficiently large waiting room to make this a very rare event.

Throughout this paper we shall limit our attention to bounded queues. In the numerical examples considered the largest value of the maximum queue length L_1 was, at most, 100.

b. Transient versus Steady-state Behavior

There are very few results on the transient behavior of queues, which are analytically explicit. Even the latter are nearly all ill-suited for direct numerical analysis. An interesting discussion of a simple transient queue and the difficulties of its computational analysis may be found in Leese and Boyd [1].

The steady-state probabilities of some simple queues are the only ones available in books written for the applied worker. Their relevance to the concrete problems is often limited; they clearly have no bearing whatsoever on the solution of unstable queues. Moreover, since the limiting process by which they are obtained has an averaging (or mixing) property, such results convey no information on the fluctuations of the queue length and the waiting times. Ignoring such fluctuations in a design may have catastrophic results.

In recent years the study of weak convergence properties of queueing processes and the resulting diffusion approximations have shed a new light on service systems of which the macroscopic time-behavior rather than the short-range fluctuations is the most important feature. This promising approach is mathematically fairly sophisticated and has not yet undergone sufficient investigation from the viewpoint of computation.

Our concern in this paper is, in a sense, with the small-scale service system whose short-range behavior is important. Therefore the models studied here are unlikely to be well approximated by a diffusion process. Nevertheless a comparison of both the direct solution and a diffusion approximation method with regards to accuracy and computing time is of interest, but will not be undertaken here.

c. Continuous versus Discrete Parameter Models

It is known that finite $M \mid G \mid 1$ and $GI \mid M \mid 1$ queues may be conveniently studied in relation to an imbedded Markov renewal process with a finite number of states. The transient behavior of such queues, as well as many related ones, may be computed in principle in terms of the successive matrix-convolution products of the transition matrix of this imbedded process. If a queue with a waiting room of size L_1 is investigated, each such a matrix-convolution product may require as many as $(L_1+1)^2$ evaluations of the convolution product of two functions. To perform this operation accurately is time-consuming, so that the computer implementation of this analysis is likely to result in considerable computing time.

In an earlier study of the single server queue in discrete time, Dafermos and Neuts [2] argued at length for the advantages of analyzing many queues in terms of a discrete time parameter. These arguments will not be repeated here. From the viewpoint of numerical analysis the most obvious advantages of a discrete time model are:

a. The ease with which one or more supplementary variables may be introduced so as to imbed the queueing process in a multivariate Markov process.

b. The fact that convolution products of sequences of numbers may be computed with much greater ease than those of functions of a continuous real variable may be evaluated accurately.

In most cases of practical interest one may discern an elementary unit of time natural to the particular queue. Many queueing analysts nevertheless insists on thinking of discrete models as approximations to continuous ones. This insistence, which may usually be traced to the prevailing attitude in applied mathematics before the advent of the computer, has some appeal for its mathematical elegance particularly where methods of analysis may be used. From a computational viewpoint, continuous parameter models are often substantially more delicate to analyze and this without yielding additional insight into the real process which is being modeled.

d. Parametric versus General Distributions

The insistence on specific parametric families of probability distributions (such as the gamma family) in stochastic models is also largely a holdover from the pre-computer era. Where a parametric assumption has an important structural consequence (e.g., the memory-less property of the negative exponential distribution) one should be very aware of this. However, in cases where the parametric assumption yields only marginal simplifications, both the theoretical analysis and the computational methods should ignore it altogether. As a case in point, the $M|E_k|1$ queue for $k > 1$ is only in some details easier to discuss than the $M|G|1$ queue. There is therefore little point in a special numerical method for the former which does not also include the latter.

In the sequel of this discussion, we shall therefore only stress the structural assumptions that are needed. Such items as service time distributions will be as general as possible.

In most specific models one may assume without loss of generality that service times and related random variables are bounded lattice random variables. This assumption, essential to our approach, is also the one which limits its range of applicability most. Our approach is not well suited to queues

in which both very short and very long jobs may arrive. Different methods of numerical analysis are needed for these. It is commonly so that a given numerical method is well suited for a certain range of problems, but fails for similar problems outside this range. The assumptions which limit the applicability of the present model are discussed in the appropriate places in the sequel.

2. THE ASSUMPTIONS OF THE MODEL

a. The Arrivals

We consider a single server queue in discrete time with a maximum queuelength L_1. The elementary time interval is chosen as our time unit. We assume that the numbers of arrivals during the successive unit time intervals are independent, identically distributed random variables. Furthermore p_ν, $\nu = 0, 1, \ldots, K$, is the probability that ν customers join the queue during a given unit time interval. ($p_0 + p_1 + \ldots + p_K = 1$). In this discussion, and in the related FORTRAN program, we shall insist that $1 \leq K < L_1$. The restriction $K < L_1$ is not essential and is usually satisfied in practice. Its removal requires minor modifications in the analysis and in the program.

We further assume that the p_ν are independent of time, but with minor obvious changes the recurrence relations are valid also for queues in which the arrival probabilities vary with time.

b. The Service Times

We assume that the service times of the successive customers are independent, identically distributed (integer-valued), random variables with values in the set $\{1, 2, \ldots, L_2\}$. We denote by r_ν, $\nu = 1, \ldots, L_2$, the probability that a customer requires ν units of service time. The values $L_1 = 100$ and $L_2 = 100$ appear to be practical upper limits to the computer implementation of the method suggested here.

Our assumption that every customer requires at least one unit of service time may easily be removed, but is satisfied in most all concrete applications. It suffices to introduce a quantity r_0 that a service time is equal to zero and to modify the recurrence relations accordingly.

As pointed out below, it is also easy to modify our analysis to include the case where the service time density depends on the time at which the service is initiated.

c. The Queue Discipline

Except in discussing the waiting times, the order of service is immaterial. The waiting times will be discussed for the first-come, first-served discipline.

To settle the issue of simultaneous arrivals and beginnings of services, we assume that all arrivals in $[n-1, n)$ are added to the queue at time $n-0$. If a service starts at time n and requires ν units of time, we shall consider it to start at time n and to end at time $n+\nu-0$.

Only as many arrivals as to maintain the queue length less than or equal to L_1 are accepted. Any excessive customers are assumed to be permanently lost.

d. The Initial Conditions

We assume that at time $n = 0$, there are i_0 customers present. $0 \leq i_0 \leq L_1$. If $i_0 \geq 1$, the customer in service at time $n = 0$ requires j_0, $1 \leq j_0 \leq L_2$ additional units of service time. We make the convention that if $i_0 = 0$, then $j_0 = 0$, and conversely.

3. THE MARKOV CHAIN

We denote by X_n the queue length at time n and by Y_n the number of additional units of service time required by the customer in service at time n. We make the convention that $X_n = 0$ if and only if $Y_n = 0$.

It follows readily from our assumptions that the bivariate sequence $\{(X_n, Y_n), n \geq 0\}$ is a Markov chain with state space consisting of the point (0, 0) and all points (i, j), $i=1, \ldots, L_1$; $j=1, \ldots, L_2$, and with initial state (i_0, j).

The transition probability matrix of the Markov chain is easily written down. We shall not do so since the recurrence relations make judicious use of its special structure.

For fixed i_0 and j_0, we define the conditional probabilities

$$P_n(i, j) = P\{X_n = i, Y_n = j \mid X_0 = i_0, Y_0 = j_0\}. \tag{1}$$

The probabilities $P_n(i, j)$ satisfy the following recurrence relations in n for all $n \geq 0$:

(2)(a) $$P_{n+1}(0, 0) = p_0[P_n(0, 0) + P_n(1, 1)],$$

(b) $$P_{n+1}(i, j) = p_0 P_n(i, j+1) + \sum_{\nu=1}^{i-1} p_{i-\nu} P_n(\nu, j+1) + r_j \left\{ p_i P_n(0, 0) + p_0 P_n(i+1, 1) + \sum_{\nu=1}^{i} p_{i-\nu+1} P_n(\nu, 1) \right\},$$

for $i=1, \ldots, K$ and $j=1, \ldots, L_2-1$.

(c) $$P_{n+1}(i, j) = p_0 P_n(i, j+1) + \sum_{\nu=i-K}^{i-1} p_{i-\nu} P_n(\nu, j+1) + r_j \left\{ p_0 P_n(i+1, 1) + \sum_{\nu=i-K+1}^{i} p_{i-\nu+1} P_n(\nu, 1) \right\},$$

for $i=K+1, \ldots, L_1-1$ and $j=1, \ldots, L_2-1$.

(d) $$P_{n+1}(L_1, j) = P_n(L_1, j+1) + \sum_{\nu=1}^{K} \left(1 - \sum_{k=0}^{\nu-1} p_k\right) P_n(L_1 - \nu, j+1) + r_j \left\{ \sum_{\nu=1}^{K} \left(1 - \sum_{k=0}^{\nu-1} p_k\right) P_n(L_1 - \nu + 1, 1) \right\}$$

for $j=1, \ldots, L_2-1$.

(e) $$P_{n+1}(i, L_2) = r_{L_2} \left\{ p_i P_n(0, 0) + p_0 P_n(i+1, 1) + \sum_{\nu=1}^{i} p_{i-\nu+1} P_n(\nu, 1) \right\},$$

for $i=1, \ldots, K$.

(f) $$P_{n+1}(i, L_2) = r_{L_2} \left\{ p_0 P_n(i+1, 1) + \sum_{\nu=i-K+1}^{i} p_{i-\nu+1} P_n(\nu, 1) \right\},$$

for $i=K+1, \ldots, L_1-1$.

(g) $$P_{n+1}(L_1, L_2) = r_{L_2} \left\{ \sum_{\nu=1}^{K} \left(1 - \sum_{k=0}^{\nu-1} p_k\right) P_n(L_1 - \nu + 1, 1) \right\}.$$

The recurrence is initialized by setting $P_0(i_0, j_0) = 1$ and $P_0(i, j) = 0$ for all other pairs (i, j). If the distribution of the random variables X_0, Y_0 is given rather than exact values, this simply amounts to a different definition of the initial array $P_0(i, j)$.

We note that the expressions in curly brackets in the formulas $(2\ b - g)$ depend on the first index only. This term corresponds to the case where the instant $n + 1$ is the beginning of a new service.

This simplifying feature is used in the organization of the computer program. If the service time distribution depends on the instant of initiation of a service, only the factor r_j in the recurrence relations are affected.

The analogous, but simpler recurrence relations for the unbounded queue were examined analytically by Dafermos and Neuts [2].

The recurrence relations (2) are well suited for iterative computation. The author organized his computation so as to have at the nth iteration only the quantities $P_n(0, 0)$ and $P_n(i, j)$, $i = 1, \ldots, L_1$; $j = 1, \ldots, L_2$ in memory.

The quantities $P_n(0, 0)$ and $P_n(i, j)$, $i = 1, \ldots, L_1$; $j = 1, \ldots, L_2$ make up the (conditional) joint density of the queue length X_n and the residual service time Y_n at time n. This joint density is not, in itself, of great practical use; however from it the distribution (or the density) of the queue length and the waiting time at time n can be easily obtained. These "derived" queue features are considered next.

4. DERIVED QUEUE FEATURES

a. The Queue Length Distribution

The probability that the queue length at time n is zero, is given by $P_n(0, 0)$. For $i = 1, \ldots, L_1$, we have

$$P\{X_n = i \mid X_0 = i_0, Y = j_0\} = \sum_{j=1}^{L_2} P_n(i, j). \tag{3}$$

Lower order moments of the queue length at time n may be obtained routinely.

b. The Waiting Time Distribution

The waiting time at time n is defined to be zero if, and only if $X_n = Y_n = 0$. For $X_n = 1$, $Y_n = j$, the waiting time equals j. For $X_n > 1$, $Y_n = j$ the waiting time is the sum of j and $X_n - 1$ independent service times. The waiting time is therefore an integer-valued random variable with value between 0 and L_1L_2.

The density WT_ν, $\nu = 0, \ldots, L_1L_2$ is obtained as follows. Clearly $WT_0 = P_n(0, 0)$. The quantities WT_n for $1 \leq \nu \leq L_1L_2$ were obtained by evaluating the following convolution-polynomial.

Let $WT(\cdot)$ denote the density $\{WT_\nu,\ 1 \leq \nu \leq L_1L_2\}$ and let $R(\cdot)$ be the service time density $\{r_j,\ 1 \leq j \leq L_2\}$. Finally let $P_n(i, \cdot)$ be the density $\{P_n(i, j),\ 1 \leq j \leq L_2\}$ for $i = 1, \ldots, L_1$, then $WT(\cdot)$ is given by

$$WT(\cdot) = P_n(1, \cdot) + P_n(2, \cdot) * R(\cdot) + \ldots + P_n(L_1, \cdot) * R^{(L_1 - 1)}(\cdot), \tag{4}$$

where $R^{(k)}(\cdot)$ is the k-fold convolution of $R(\cdot)$ with itself.

The density $WT(\cdot)$ may be computed for each n by one of two procedures. Either the convolutions $R^{(k)}(\cdot)$, $k = 1, \ldots, L_1 - 1$ may be computed once and for all and only the convolutions with the arrays $P_n(i, \cdot)$ need to be computed at those time points n at which the density of the waiting time is

desired. This procedure results in a substantial increase in the central memory storage required by the program.

Alternatively, the density $WT(\cdot)$ may be computed by a convolution analogue of *Horner's algorithm* for the numerical evaluation of ordinary polynomials. This procedure consists in the successive evaluation of the sequences

$$WT^{(1)}(\cdot) = P_n(L_1, \cdot), \tag{5}$$

$$WT^{(k)}(\cdot) = WT^{(k-1)}(\cdot) * R(\cdot) + P_n(L_1 - k + 1, \cdot),$$

for $k = 1, \ldots, L_1$. The sequence $WT^{(L_1)}(\cdot)$ is the desired sequence $WT(\cdot)$.

The latter procedure does not result in the use of core storage for intermediate quantities. For purposes of comparison and as a guard against rounding errors, both procedures were programmed and tested on large scale examples. The Horner convolution algorithm performed slightly better in all examples and no evidence of rounding errors were found. Its much smaller core requirements make it the more efficient of the two procedures.

5. REPORT ON COMPUTATIONAL TRIALS

A FORTRAN IV program was written by the author and tested on a variety of cases on the CDC 6500 at Purdue University. Even in the largest examples no evidence of rounding errors was found, even though all probabilities were printed to five decimal places and all computations were performed in single precision.

The full output consisted, in addition to the summary of the input data, of the following:

(a) the mean of the queue length,

(b) the cumulative distribution of the queue length,

(c) the mean waiting time,

(d) the cumulative distribution of the waiting time,

(e) the joint density of the queue length X_n and the residual service time Y_n.

All these for all values of n up to some upper value to be specified.

Since the full output is very voluminous and contains much more information than may be needed in the analysis of a given queue, options were written into the program which permit the deletion of the items (c), (d), or (e) from the printed output. A further option was created which permits the computation of the waiting time distribution to be performed at certain specified time points only.

No systematic study of the processing time was made. For smaller examples the computation times were generally below 10 seconds. The following computation times for some larger examples are given for purposes of illustration only.

$L_1 = 30$	(max. queue length)
$L_2 = 6$	(max. duration of one service)
$K = 2$	(max. number of arrivals)
$NNN = 250$	(number of time points computed)

In the following CP is the processing time in seconds and LL is the number of lines of output, including the program listing.

a. full printed output

$$CP = 215.159 \qquad LL = 19265$$

b. the queue length and waiting time distributions only.

$$CP = 157.860 \qquad LL = 10656$$

c. the queue length distribution for all time points and the waiting time distribution only for time points which are multiples of 25.

$$CP = 23.763 \qquad LL = 4916$$

The processing time for a large scale iteration of the type needed here can be substantially decreased by writing a program in an assembly language, rather than in FORTRAN. A reduction by 50 or 60 percent can realistically be expected.

For large problems, the computation of the waiting time distribution is by far the most time consuming part of the algorithm. For very stable queues, we recommend neglecting the higher order terms in the convolution polynomial (4). These terms contribute very little, except to the extreme upper tail of the distribution, but add very considerably to the processing time. For queues with a very rapid build-up, some savings may be accomplished by neglecting lower order terms in (4), but this is less significant. For queues which are near-critical, the waiting time algorithm ceases to be practical for queues with $L_1 > 200$, because of the large numbers of operations involved. The problem of finding good numerical approximations to the distribution of the waiting time in this case is challenging and needs further investigation.

6. CONCLUSIONS

By the use of only the most elementary structural properties, we have shown that the transient behavior of a substantial class of single server queues may be analyzed numerically. The approach presented here should prove itself to be useful in studying the build-up of unstable queues and the fluctuations of queues at traffic lights, highway merging ramps, service counters in public offices and retail outlets and many others.

This approach is well suited for many queueing processes which do not lend themselves to diffusion approximation methods. The amount of computing time used in typical examples also indicates a very substantial saving over that needed to analyze similar models by simulation methods.

Further work is currently being done to extend the applicability of this approach to longer queues and to much longer time periods. This extension, however, requires the use of mathematically more sophisticated properties of the queueing process.

A version of this paper, containing a program listing and a specific numerical example, is available as a technical report. It may be obtained from the author upon request, by writing to Professor Marcel F. Neuts, Department of Statistics, Purdue University, West Lafayette, IN. 47907.

BIBLIOGRAPHY

[1] Leese, E. L. and D. W. Boyd, "Numerical Methods of Determining the Transient Behavior of Queues with Variable Arrival Rates," J. Canadian Operations Res. Soc. *4*, 1–13 (1966).

[2] Dafermos, S. and M. F. Neuts, "A Single Server in Discrete Time," Cahiers du Centre de Recherche Opérationnelle, *13*, 23–40 (1971).

THE SINGLE SERVER QUEUE IN DISCRETE TIME—NUMERICAL ANALYSIS II

Eugene M. Klimko and Marcel F. Neuts*

Purdue University

ABSTRACT

This paper deals with the numerical problems arising in the computation of higher order moments of the busy period for certain classical queues of the $M|G|1$ type, both in discrete and in continuous time.

The classical functional equation for the moment generating function of the busy period is used. The higher order derivatives at zero of the moment generating function are computed by repeated use of the classical differentiation formula of Fáa di Bruno. Moments of order up to fifty may be computed in this manner.

A variety of computational aspects of Fáa di Bruno's formula, which may be of use in other areas of application, are also discussed in detail.

1. INTRODUCTION

This paper deals with the numerical computation of moments of high order of the busy period for several fundamental queueing models. We shall specifically consider the following queueing processes.

The Models

For brevity we refer to the two models as (a) *the continuous model* and (b) *the discrete model*, respectively.

(*a*) *The continuous model*—This is the classical $M^{(X)}|G|1$ queue in which customers arrive in groups of random size in epochs which form a homogeneous Poisson process of rate λ. They are served, one at a time, and their successive service times are independent and identically distributed random variables with the common distribution $H(\cdot)$. We further assume that the successive group sizes are independent, identically distributed random variables. The probability that an arrival consists of a group of exactly k customers is denoted by p_k, where $p_0=0$ and $\sum_{k=1}^{\infty} p_k=1$.

The nth moments of the distribution $H(\cdot)$ and of the discrete density $\{p_k\}$ are denoted by α_n and η_n, respectively. We assume throughout this paper that all moments considered are finite.

The busy period of this queue is defined as the length of time until a server, starting with a single customer, becomes idle for the first time. For every $x \geq 0$, $G(x)$ denotes the probability that the duration of the busy period does not exceed x. The function $G(\cdot)$ is then a (possibly defective) probability distribution. The following theorems are well-known, but are stated here for completeness.

*The research of this author was supported by the National Science Foundation grant GP 28650.

THEOREM 1: For every $x \geq 0$, the function $G(\cdot)$ satisfies the nonlinear integral equation

$$G(x) = \sum_{\nu=0}^{\infty} \int_0^x e^{-\lambda y} \frac{(\lambda y)^\nu}{\nu!} \sum_{r=\nu}^{\infty} p_r^{(\nu)} G^{(r)}(x-y)\, dH(y), \tag{1}$$

in which $G^{(r)}(\cdot)$ is the r-fold convolution of $G(\cdot)$ with itself and $\{p_r^{(\nu)}\}$ is the ν-fold convolution of the discrete density $\{p_r\}$ with itself.

PROOF: By an application of the law of total probability. Let there be exactly ν arrivals during the service of the first customer and let there be exactly $r \geq \nu$ customers in the ν arriving groups. Each of these r customers can be considered as the initial one of r independent busy periods. If the first service has a duration y, then the r busy periods so generated can last, at most for a length of time $x-y$. This argument is due to L. Takàcs and may be found in more detail in [8, p. 47].

Introducing *the moment generating functions* $\gamma(s)$, $h(s)$, and $\theta(s)$ of $G(\cdot)$, $H(\cdot)$, and $\{p_k\}$, i.e.,

$$\gamma(s) = \int_0^\infty e^{sx} dG(x), \qquad h(s) = \int_0^\infty e^{sx} dH(x), \tag{2}$$

$$\theta(s) = \sum_{k=1}^{\infty} p_k e^{ks},$$

for $Re\ s \leq 0$, the integral Equation (1) may be equivalently written as

$$\gamma(s) = h[s - \lambda + \lambda\theta[\log \gamma(s)]], \tag{3}$$

where *log* denotes the principal branch of the logarithmic function.

THEOREM 2: For every s with $Re\ s < 0$, the functional Equation (3) has a unique solution in the unit disk. This solution $\gamma(s)$ is analytic in the halfplane $Re\ s < 0$ and continuous on the boundary. Moreover $\gamma(\cdot)$ is the moment generating function of a (possibly defective) probability distribution $G(\cdot)$.

PROOF: This result, which is usually proved by applying Rouché's theorem, is a classical theorem in the theory of queues.

THEOREM 3: The probability distribution $G(\cdot)$ is proper, i.e., $G(+\infty) = 1$, if and only if

$$\lambda \eta_1 \alpha_1 \leq 1. \tag{4}$$

If equality holds, the first moment of the busy period is infinite.

PROOF: See e.g., Takàcs [7].

REMARK: Throughout this paper we consider only *stable* queues, i.e., we assume henceforth that

$$\lambda \eta_1 \alpha_1 < 1. \tag{5}$$

(*b*) *The discrete model* — As discussed by S. Dafermos and M. F. Neuts [1], there is a considerable advantage, particularly from a computational viewpoint, in considering the discrete-time analogue of the model (*a*). The time variable is now discrete. The numbers of customers arriving during successive

unit time intervals are independent, identically distributed integer-valued random variables with discrete density $\{p_k'\}$. p_0' is now (usually) positive and is the probability that no customer arrives during a unit of time.

Customers are served singly and the probability that a customer requires ν units of service time is denoted by r_ν for $\nu \geq 1$. The usual independence assumptions are made.

The busy period is again the length of time until a server, starting with one customer, becomes idle for the first time. We shall denote by β_n, $n \geq 1$, the probability that the busy period lasts for exactly n units of time. $\{\beta_n\}$ is a (possibly defective) discrete probability density.

THEOREM 4: The density $\{\beta_n\}$ satisfies the nonlinear recurrence relation

$$\beta_n = \sum_{\nu=1}^{n} r_\nu \sum_{k=0}^{\infty} p_k'^{(\nu)} \beta_{n-\nu}^{(k)}, \tag{6}$$

for $n \geq 1$. In Equation (6), $\{p_k'^{(\nu)}\}$ and $\{\beta_k^{(\nu)}\}$ are the ν-fold convolutions of the densities $\{p_k'\}$ and $\{\beta_k\}$.

PROOF: By an application of the law of total probability as in the proof of Theorem 1, (6) follows.

We now introduce the moment generating functions

$$B(s) = \sum_{n=1}^{\infty} \beta_n e^{ns}, \quad R(s) = \sum_{n=1}^{\infty} r_n e^{ns}, \quad P(s) = \sum_{n=0}^{\infty} p_n' e^{ns}, \tag{7}$$

for $Re\, s \leq 0$ and obtain by using Equation (6) that

$$\begin{aligned} B(s) &= \sum_{n=1}^{\infty} \beta_n e^{ns} = \sum_{\nu=1}^{\infty} \sum_{n=\nu}^{\infty} r_n e^{ns} \sum_{k=0}^{\infty} p_k'^{(\nu)} \beta_{n-\nu}^{(k)} \\ &= \sum_{\nu=1}^{\infty} r_\nu e^{\nu s} \sum_{\tau=0}^{\infty} e^{\tau s} \sum_{k=0}^{\infty} p_k'^{(\nu)} \beta_\tau^{(k)} \\ &= \sum_{\nu=1}^{\infty} r_\nu e^{\nu s} \sum_{k=0}^{\infty} p_k'^{(\nu)} B^k(s) = \sum_{\nu=1}^{\infty} r_\nu e^{\nu s} P^\nu [\log B(s)] \\ &= R\,[s + \log P\,[\log B(s)]], \end{aligned} \tag{8}$$

for $Re\, s < 0$. In both cases *log* is the principal branch of the logarithmic function.

THEOREM 5: The functional Equation (8) has a unique solution inside the unit disk for every s with $Re\, s < 0$. The solution $B(s)$ is analytic inside the region $Re\, s < 0$ and continuous on the boundary and the function $B(\cdot)$ is the moment generating function of a discrete density $\{\beta_n\}$. The density $\{\beta_n\}$ is proper, i.e., $\sum_{n=1}^{\infty} \beta_n = 1$, if and only if

$$\eta_1 \alpha_1 \leq 1, \tag{9}$$

where $\alpha_1 = R'(0-)$ is the mean service time and $\eta_1 = P'(0-)$ is the mean number of arrivals per unit of time. When equality holds in (9), the *mean of the busy period* is infinite.

PROOF: This theorem is proved in its generating function version in S. Dafermos and M. F. Neuts [1].

REMARK: In the remainder of this paper we shall again restrict our attention to the case of *stable* queues and we assume therefore that $\eta_1\alpha_1 < 1$.

2. STATEMENT AND SIGNIFICANCE OF THE PROBLEM

The moments of the busy period may be obtained by successive differention with respect to s in Equation (3) for the continuous case and in Equation (8) for the discrete case. By taking the limit as $s \to 0-$ and by appealing to Abel's theorem, we obtain a relation expressing the nth moment in terms of the moments of orders one up to $n-1$. However, in view of the multiple functional compositions which occur on the right-hand sides of Equations (3) and (8), this recurrence relation soon becomes very unwieldy and its value for numerical calculations is far from obvious.

The purpose of our discussion is to show that, on the contrary, the recurrence relations generated by successive differentiations are practical for the evaluation of moments *up to order 50*, approximately. Moreover the application of this technique is not solely limited to problems of queueing theory, which are discussed here in some detail. Functional compositions of generating functions occur in many problems of probability theory, for example in applications involving mixtures or compound distributions. The algorithm is also useful in the numerical computation of moments in such cases. It should also find application in procedures for numerical differentiation.

Knowledge of the higher moments of the busy period is useful, in particular for the following reason. Equations (1) and (6) are both of the general form

$$G = \sum_{n=0}^{\infty} A_n * G^{(n)}. \tag{10}$$

The former is now a nonlinear integral equation and the latter a nonlinear difference equation. Both can be shown to be well-suited for numerical solution by successive substitution methods. We shall report in a subsequent paper on procedures for selecting a good starting solution and on the computer implementation of the method of iterative solution.

However, a very important step in this algorithm is the selection of a practical upper bound for the support of the distribution $G(\cdot)$ or of the density $\{\beta_n\}$. Explicitly, one needs to find a quantity A such that for a given error term ϵ, (say 10^{-4}), either

$$1 - G(A) \leq \epsilon, \qquad \text{or} \sum_{\nu=A+1}^{\infty} \beta_\nu \leq \epsilon. \tag{11}$$

An upper bound for A is given as $[\epsilon^{-1}EX^n]^{1/n}$ by Markov's inequality, in terms of the nth moment. In [6], it was shown that for sufficiently small ϵ, this upper bound is logarithmically convex in n. By successive computation of higher moments, we may therefore obtain the tightest Markov upper bound in many practical caces. When many moments are known, tighter upper bounds may also be used, but these are not as readily computed as the Markov bound.

3. THE RECURRENCE RELATION

The functional equations, Equations (3) and (8), are very similar in nature. We shall give a detailed discussion of Equation (3) and indicate the appropriate changes for application of our results to Equa-

tion (8). The nth moment of the busy period will be denoted by g_n. Clearly

$$g_n = \gamma^{(n)}(0-), \qquad \text{for } n \geq 1. \tag{12}$$

This leads for $n=1$ and $n=2$ to the well-known formulas

$$g_1 = \alpha_1(1-\lambda\eta_1\alpha_1)^{-1}, \tag{13}$$

$$g_2 = [\alpha_2 + \lambda(\eta_2-\eta_1)\alpha_1^3](1-\lambda\eta_1\alpha_1)^{-3}.$$

In order to express the nth derivative we appeal to the classical formula of Fáa di Bruno [2], which expresses the nth derivative of a composite function $f[\phi(x)]$ as follows.

Fáa di Bruno's Formula

Assuming the existence of all the derivatives involved, we have that

$$\left\{\frac{d^n}{dx^n} f[\phi(x)]\right\}_{x=0} = \sum_{r=1}^{n} \left\{\frac{d^r}{dy^r} f(y)\right\}_{y=\phi(0)} \sum_{\substack{j_1+\ldots+j_n=r \\ j_1+2j_2+\ldots+nj_n=n \\ j_i\geq 0,\ldots,j_n\geq 0}} \frac{n!}{j_1!j_2!\ldots j_n!} \tag{14}$$

$$\times \left(\frac{\phi^{(1)}(0)}{1!}\right)^{j_1} \left(\frac{\phi^{(2)}(0)}{2!}\right)^{j_2} \ldots \left(\frac{\phi^{(n)}(0)}{n!}\right)^{j_n}.$$

In general, assuming the existence of all the derivatives involved, the nth moment satisfies

$$g_n = \left\{\frac{d^n}{ds^n} h[s-\lambda+\lambda\theta[\log\gamma(s)]]\right\}_{s=0-} = \sum_{r=1}^{n} \alpha_r Y_{nr}^{(1)}, \tag{15}$$

where

$$Y_{nr}^{(1)} = \sum_{\substack{j_1+\ldots+j_n=r \\ j_1+2j_2+\ldots+nj_n=n \\ j_1\geq 0,\ldots,j_n\geq 0}} \frac{n!}{j_1!\ldots j_n!} A_1^{j_1}A_2^{j_2}\ldots A_n^{j_n}. \tag{16}$$

The quantities $A_1, \ldots, A_n$ are given by

$$A_1 = 1+\lambda\theta'(0)\gamma'(0) = 1+\lambda\eta_1 g_1 = (1-\lambda\eta_1\alpha_1)^{-1}, \tag{17}$$

and for $2 \leq \nu \leq n$

$$A_\nu = \frac{\lambda}{\nu!}\left\{\frac{d^\nu}{dx^\nu}\theta[\log\gamma(x)]\right\}_{x=0}.$$

The latter derivatives are given by

$$\left\{\frac{d^\nu}{dx^\nu}[\log\gamma(x)]\right\}_{x=0}=\sum_{m=1}^{\nu}\eta_m Y_{\nu m}^{(2)}, \tag{18}$$

for $2\leq\nu\leq n$, where

$$Y_{\nu m}^{(2)}=\sum_{\substack{i_1+i_2+\ldots+i_\nu=m\\ i_1+2i_2+\ldots+\nu i_\nu=\nu\\ i_1\geq 0,\ldots,i_\nu\geq 0}}\frac{\nu!}{i_1!i_2!\ldots i_\nu!}B_1^{i_1}B_2^{i_2}\ldots B_\nu^{i_\nu}. \tag{19}$$

The quantities B_ρ are given by

$$B_1=\gamma'(0)=g_1=\alpha_1(1-\lambda\eta_1\alpha_1)^{-1}, \tag{20}$$

and for $2\leq\rho\leq\nu\leq n$

$$B_\rho=\frac{1}{\rho!}\left\{\frac{d^\rho}{dx^\rho}\log\gamma(x)\right\}_{x=0}.$$

The latter derivatives, in turn, are given by

$$\left\{\frac{d^\rho}{dx^\rho}\log\gamma(x)\right\}_{x=0}=\sum_{h=1}^{\rho}(-1)^{h+1}(h-1)!Y_{\rho h}^{(3)}, \tag{21}$$

where

$$Y_{\rho h}^{(3)}=\sum_{\substack{\tau_1+\ldots+\tau_\rho=h\\ \tau_1+2\tau_2+\ldots+\rho\tau_\rho=\rho\\ \tau_i\geq 0,\ldots,\tau_\rho\geq 0}}\frac{\rho!}{\tau_1!\tau_2!\ldots\tau_\rho!}\left(\frac{g_1}{1!}\right)^{\tau_1}\left(\frac{g_2}{2!}\right)^{\tau_2}\ldots\left(\frac{g_\rho}{\rho!}\right)^{\tau_\rho}. \tag{22}$$

The formulas (15)–(22) may be combined to express g_n as a (complicated) polynomial in $g_1, \ldots, g_{n-1}, g_n$. A notable simplication occurs, however, if we observe that g_n occurs *only once* among the terms on the right. By examining the successive conditions on the indices in the three applications of Fáa di Bruno's formula, we find that the *only* term containing g_n which appears on the right hand side is

$$(\lambda\eta_1\alpha_1)g_n.$$

It follows that

$$(1-\lambda\eta_1\alpha_1)g_n=\sum_{r=2}^{n}Y_{nr}^{(1)}\alpha_r+\lambda\alpha_1\sum_{m=2}^{n}Y_{nm}^{(2)}\eta_m+\lambda\alpha_1\eta_1\sum_{h=2}^{n}(-1)^{h+1}(h-1)!Y_{nh}^{(3)}. \tag{23}$$

The expression on the right is a polynomial of degree n in $g_1, g_2, \ldots g_{n-1}$. By application of the formulas (16)–(23), we compute the higher moments of the busy period recursively. It is clear that the practical

limitations on this method depend mainly on the growth of the number of terms appearing on the right hand side in Fáa di Bruno's formula. This matter is discussed below, but first we indicate the modifications necessary for the discrete case and for some particular cases.

The Case of Single Arrivals

When customers arrive singly, we have $\theta(s)=e^s$, so that Equation (3) reduces to

$$\gamma(s)=h[s-\lambda+\lambda\gamma(s)]. \tag{24}$$

This case requires only *two* applications of Fáa di Bruno's formula.

The Discrete Case

In general, the discrete case requires four applications of Fáa di Bruno's formula. The basic formulas in this case are

$$g_1=\alpha_1(1-\lambda_1\alpha_1)^{-1}, \tag{25}$$

and

$$g_2=[\alpha_2+\alpha_1(\eta_2-\eta_1^2-\eta_1)](1-\eta_1\alpha_1)^{-3}$$

$$g_n=\sum_{r=1}^{n}\alpha_r Z_{nr}^{(1)}, \qquad \text{for } n\geq 3, \tag{26}$$

where

$$Z_{nr}^{(1)}=\sum_{\substack{j_1+\ldots+j_n=r\\ j_1+2j_2+\ldots+nj_n=n\\ j_i\geq 0,\ldots,j_n\geq 0}}\frac{n!}{j_1!\ldots j_n!}A_1^{j_1}\ldots A_n^{j_n} \tag{27}$$

for $1\leq r\leq n$.

$$A_1=(1-\eta_1\alpha_1)^{-1}. \tag{28}$$

$$A_\nu=\frac{1}{\nu!}\left\{\frac{d^\nu}{dx^\nu}\log P[\log B(x)]\right\}_{x=0}, \tag{29}$$

for $2\leq\nu\leq n$. Furthermore

$$\left\{\frac{d^\nu}{dx^\nu}\log P[\log B(x)]\right\}_{x=0}=\sum_{m=1}^{\nu}(-1)^{m+1}(m-1)!\ Z_{\nu m}^{(2)} \tag{30}$$

where

$$Z_{\nu m}^{(2)} = \sum_{\substack{j_1+\ldots+j_\nu=m \\ j_1+2j_2+\ldots+\nu j_\nu=\nu \\ j_1\geq 0,\ldots,j_\nu\geq 0}} \frac{\nu!}{j_1!\ \ldots\ j_\nu!} C_1^{j_1}\ \ldots\ C_\nu^{j_\nu}, \tag{31}$$

where

$$C_1 = \eta_1 g_1 = \alpha_1\eta_1(1-\alpha_1\eta_1)^{-1}, \tag{32}$$

and

$$C_\rho = \frac{1}{\rho!}\left\{\frac{d^\rho}{dx^\rho} P[\log B(x)]\right\}_{x=0}, \qquad \text{for } 2\leq\rho\leq\nu\leq n. \tag{33}$$

The latter derivative is given by

$$\left\{\frac{d^\rho}{dx^\rho} P[\log B(x)]\right\}_{x=0} = \sum_{\tau=1}^{\rho} \eta_\tau Z_{\rho\tau}^{(3)}, \tag{34}$$

where

$$Z_{\rho\tau}^{(3)} = \sum_{\substack{j_1\ldots+j\rho=\tau \\ j_1+2j_2+\ldots+\rho j_\rho=\rho \\ j_1\geq 0,\ldots,j_\rho\geq 0}} \frac{\rho!}{j_1!\ \ldots\ j_\rho!} E_1^{j_1}\ \ldots\ E_\rho^{j_\rho}. \tag{35}$$

the quantities E_q, for $1\leq q\leq\tau\leq\nu\leq n$ are given by

$$E_q = \frac{1}{q!}\left\{\frac{d^q}{dx^q}\log B(x)\right\}_{x=0}, \tag{36}$$

where

$$\left\{\frac{d^q}{dx^q}\log B(x)\right\}_{x=0} = \sum_{u=1}^{q} (-1)^{u+1}(u-1)!\, Z_{qu}^{(4)}, \tag{37}$$

and

$$Z_{qu}^{(4)} = \sum_{\substack{j_1+\ldots+j_q=u \\ j_1+2j_2+\ldots+qj_q=q \\ j_1\geq 0,\ldots,j_q\geq 0}} \frac{u!}{j_1!\ \ldots\ j_q!}\left(\frac{g_1}{1!}\right)^{j_1}\ \ldots\ \left(\frac{g_q}{q!}\right)^{j_q}, \tag{38}$$

for $u=1, \ldots, q$.

A similar observation as in the continuous case can be made here. By examination of the conditions on the indices of the Fáa di Bruno formulas, we note that the only term involving g_n which appears on the right-hand side is $\alpha_1\eta_1 g_n$. It follows that the quantity $(1-\alpha_1\eta_1)\ g_n$ may be expressed as a polynomial in $g_1, \ldots, g_{n-1}$.

REMARK: It is often convenient to choose the mean service time α_1 as a new unit of time. The normalized nth moment $\tilde{g}_n$ corresponding to $\alpha_1 = 1$ is related to the one corresponding to a general value of α_1 by the formula

$$g_n = \tilde{g}_n \alpha_1^n. \tag{39}$$

4. THE MOMENTS FOR THE BUSY PERIOD OF THE $M|M|1$ QUEUE.

Explicit expressions for the higher moments of the busy period are available in rare cases only. As is often the case, the $M|M|1$ queue leads to tractable expressions. The moments of the $M|M|1$ queue were used to verify the accuracy and the correctness of our general computer routines.

In this section, we study the moments $\tilde{g}_n$ of the busy period for the $M|M|1$ queue, first with single arrivals and next for the case of geometrically distributed bunch sizes.

a. Single Arrivals

The moment generating function for the busy period of the queue with Poisson arrivals of rate ρ and with mean service time $\mu = 1$ is given by

$$\gamma(s) = \frac{1+\rho-s}{2\rho} - \frac{1-\rho}{2\rho}\left[1 - s\,(1-\sqrt{\rho})^{-2}\right]^{1/2}\left[1 - s\,(1+\sqrt{\rho})^{-2}\right]^{1/2}. \tag{40}$$

After a routine series expansion, using the binomial series, one obtains

$$\gamma(s) = 1 + s(1-\rho)^{-1} + \frac{1-\rho}{2\rho}\sum_{n=2}^{\infty}(-1)^{n+1}s^n \sum_{\nu=0}^{n}\binom{1/2}{\nu}\binom{1/2}{n-\nu}(1-\sqrt{\rho})^{-2\nu}(1+\sqrt{\rho})^{2\nu-2n}. \tag{41}$$

The moments $\tilde{g}_n$ of the busy period are therefore given by

$$\tilde{g}_1 = (1-\rho)^{-1} \tag{42}$$

and

$$\tilde{g}_n = n!\,\frac{(1-\rho)}{2\rho}(-1)^{n+1}\sum_{\nu=0}^{n}\binom{1/2}{\nu}\binom{1/2}{n-\nu}(1-\sqrt{\rho})^{-2\nu}(1+\sqrt{\rho})^{2\nu-2n}, \tag{43}$$

for $n \geq 2$.

Simplified expressions for $2 \leq n \leq 5$ are

$$\tilde{g}_2 = 2(1-\rho)^{-3} \tag{44}$$

$$\tilde{g}_3 = 3!(1-\rho)^{-5}(1+\rho),$$

$$\tilde{g}_4 = 4!(1-\rho)^{-7}(1+3\rho+\rho^2),$$

$$\tilde{g}_5 = 5!(1-\rho)^{-9}(1+\rho)(1+5\rho+\rho^2).$$

In order to evaluate $\tilde{g}_n$ numerically we wrote $\tilde{g}_n$ in the form

$$\tilde{g}_n = A'_{n0} - \sum_{\nu=1}^{n-1} A'_{n\nu} + A'_{nn}, \tag{45}$$

where

$$A'_{n0} = n! \frac{1-\rho}{2\rho} (-1)^{n+1} \binom{1/2}{n} (1+\sqrt{\rho})^{-2n}, \tag{46}$$

$$A'_{nn} = n! \frac{1-\rho}{2\rho} (-1)^{n+1} \binom{1/2}{n} (1-\sqrt{\rho})^{-2n},$$

and

$$A'_{n\nu} = n! \frac{1-\rho}{2\rho} (-1)^{n} \binom{1/2}{\nu} \binom{1/2}{n-\nu} (1-\sqrt{\rho})^{-2}(1+\sqrt{\rho})^{2\nu-2n},$$

for $\nu = 1, \ldots, n-1$.

We further noted the following recurrence relations

$$A'_{n0} = A'_{n-1,0}(n-3/2)(1+\sqrt{\rho})^{-2}, \tag{47}$$

$$A'_{n,n} = A'_{n-1,n-1}(n-3/2)(1-\sqrt{\rho})^{-2},$$

$$A'_{n,\nu} = A'_{n-1,\nu} n \left[1 - \frac{3}{2}\frac{1}{n-\nu}\right](1+\sqrt{\rho})^{-2}, \qquad \text{for } \nu = 1, \ldots, n-2,$$

$$A'_{n,n-1} = A'_{n-1,n-1} \frac{n}{2} (1+\sqrt{\rho})^{-2}.$$

For each n, the quantities $A'_{n,\nu}$, $0 \leq \nu \leq n$ are readily computed in terms of the corresponding quantities for $n-1$. The accuracy of the single precision calculation was checked by comparison with a double precision routine and, for the values computed, agreement was found for the first 11 significant digits. Moments up to order 80 were computed in this manner.

b. Group Arrivals

It was further desirable to have explicit expressions for a queue with group arrivals.

Explicit expressions for the moments $\tilde{g}_n^*$ for the queue with exponential service times and geometric bunch size distribution were obtained. For this queue, the moment generating function, $\gamma(s)$, satisfies the quadratic equation

$$(\lambda+q-qs)\gamma^2(s)-(\lambda+q+1-s)\gamma(s)+1=0, \tag{48}$$

where p is the parameter of the geometric bunch size distribution, $q=1-p$ and λ is the arrival rate of the groups of customers. The solution in the unit disk of this equation is given by

(49)

$$\gamma(s)=\frac{(\lambda+q+1)-s}{2(\lambda+q)}\left(1-\frac{qs}{\lambda+q}\right)^{-1}+\frac{\lambda-p}{2(\lambda+q)}\left(1-\frac{qs}{\lambda+q}\right)^{-1}\left[1-\frac{s}{(\sqrt{\lambda}+\sqrt{p})^2}\right]^{1/2}\left[1-\frac{s}{(\sqrt{\lambda}-\sqrt{p})^2}\right]^{1/2}.$$

Using the negative binomial and the geometric series and performing several multiplications of series, we found the expansion

(50)

$$\gamma(s)=\frac{\lambda+q+1}{2(\lambda+q)}\left\{1+\frac{q^2-\lambda p}{q(\lambda+q+1)}\sum_{r=1}^{\infty}\left(\frac{q}{\lambda+q}\right)^r s^r\right\}$$

$$+\sum_{r=0}^{\infty}s^r\sum_{k=0}^{r}\left(\frac{q}{\lambda+r}\right)^{r-k}(-1)^k\sum_{\nu=0}^{k}\binom{1/2}{\nu}\binom{1/2}{k-\nu}(\sqrt{\lambda}+\sqrt{p})^{-2\nu}(\sqrt{\lambda}-\sqrt{p})^{2\nu-2k}.$$

It is known that this queue is stable if and only if $\lambda \leq p$. If we denote $\frac{\lambda}{p}$ by θ, then the mean of the busy period is given by

$$\tilde{g}_1^*=(1-\theta)^{-1}. \tag{51}$$

The higher moments $\tilde{g}_n^*$ of the busy period are simply related to the moments $\tilde{g}_k(\theta)$, $k \geq 1$, of the $M|M|1$ queue with *single* arrivals and traffic intensity θ. Upon examination of the coefficient of s^n in the series expansion (50) and by using formula (43), we obtain

$$\tilde{g}_n^*=n!\left(\frac{q}{\lambda+q}\right)^n+\frac{\lambda}{\lambda+q}n!\sum_{k=1}^{n}\left(\frac{q}{\lambda+q}\right)^{n-k}\frac{1}{p^k k!}\tilde{g}_k(\theta) \tag{52}$$

for $n \geq 2$. Formula (52) indicates clearly how the quantities $\tilde{g}_n^*$ can be computed by a routine modification of the algorithm used for the evaluation of the moments $\tilde{g}_n(\theta)$.

5. NUMERICAL ASPECTS OF FÁA DI BRUNO'S FORMULA

a. The Number of Terms

In order to study the number of terms appearing on the right-hand side in Fáa di Bruno's formula, let $\Psi(n, r, s)$ be the numer of n-tuples $(j_1, \ldots, j_n)$ of nonnegative integers satisfying

$$j_1+j_2+\ldots+j_n=r, \qquad j_1+2j_2+\ldots+nj_n=s, \qquad \text{for } s \geq r. \tag{53}$$

The quantities $\Psi(n, r, s)$ satisfy the recurrence relations

$$\Psi(n, r, s) = \Psi(n-1, r, s) + \sum_{j=1}^{j^*} \Psi(n-1, r-j, s-nj), \tag{54}$$

where

$$j^* = \min\left(r, \frac{s-r}{n-1}\right), \qquad \text{for } n \geq 2. \tag{55}$$

If $j^* = 0$, then the latter summation is vacuous. The recurrence is initialized by $\Psi(1, 1, 1) = 1$.

The number of terms in Fáa di Bruno's formula is then given by

$$\sum_{r=1}^{n} \Psi(n, r, n) = \Psi^*(n). \tag{56}$$

Using (54), we computed $\Psi^*(n)$ for n up to 100. The growth of $\Psi^*(n)$ is quite slow up to $n = 30$, is moderate between $n = 31$ and $n = 40$, and is fast from $n = 40$ on. The terms slow, moderate, and fast are, or course, strictly qualitative. We note the following values of $\Psi^*(n)$.

$$\begin{aligned} \Psi^*(10) &= \quad 42, \\ \Psi^*(20) &= \quad 627, \\ \Psi^*(30) &= \quad 5{,}604, \\ \Psi^*(40) &= \; 37{,}338, \\ \Psi^*(50) &= 204{,}226. \end{aligned} \tag{57}$$

Again in qualitative terms, and in relation to the speed of present-day computers, we observe that Fáa di Bruno's formula is *readily applicable* for $n \leq 30$, whereas $n = 50$ is very much a practical upper bound for its computational use.

b. The number of Nonzero Indices

Computationally it is advantageous to use only the nonzero indices in Fáa di Bruno's formula. We therefore investigated the maximum number $\chi(n)$ of positive integers j_ν occurring in an n-tuple $(j_1, \ldots, j_n)$ nonnegative integers, satisfying

$$j_1 + 2j_2 + \ldots + nj_n = n. \tag{58}$$

The practical value of Fáa di Bruno's formula is greatly enhanced by the *slow growth* of $\chi(n)$. A numerical investigation of $\chi(n)$ for $n \leq 50$ showed that

$$\begin{aligned} \chi(n) = 2, &\qquad \text{for } 3 \leq n \leq 5, \\ 3, &\qquad \text{for } 6 \leq n \leq 9, \end{aligned} \tag{59}$$

$$\begin{aligned} 4, &\quad \text{for } 10 \leq n \leq 14, \\ 5, &\quad \text{for } 15 \leq n \leq 20, \\ 6, &\quad \text{for } 21 \leq n \leq 27, \\ 7, &\quad \text{for } 28 \leq n \leq 35, \\ 8, &\quad \text{for } 36 \leq n \leq 44, \\ 9, &\quad \text{for } 45 \leq n \leq 50. \end{aligned}$$

Knowledge of the values of $\chi(n)$ is very important for designing computer programs for calculating the busy period moments.

6. COMPUTER PROGRAM CONSIDERATIONS AND RESULTS

The authors have designed a system of computer routines to implement the recurrence relations for $\tilde{g}_n$, derived in section 3. The main problem lies in the large number of terms occuring in the summations used to evaluate the coefficients Y_{nr} in Fáa di Bruno's formula. The considerations of section 5 show that the volume of computation required for the higher moments, dictates efficient program organization.

Since Fáa di Bruno's formula is applied several times, it is advantageous to have a list of the indices $(j_1, j_1, \ldots j_n)$ stored in memory rather than to generate them repeatedly. Therefore the computer routines were divided into two sets.

The *first* set generates a list of the nonzero indices $(j_1, j_2, \ldots, j_n)$ of summation in Fáa di Bruno's formula. This list may be stored once and for all on a magnetic tape or in a disk storage device. The indices are subsequently read from this storage medium. The *second* set of programs reads the indices from their storage medium and performs the computations appropriate for the main recurrence relation.

Not all indices are read in from the storage medium at once. In order to conserve memory space, a first set of indices is read and is used to compute all the quantities Y_{nr} in which they occur. A new set of indices, corresponding to higher values of n is then read in and is stored in the same locations as those of the previous set.

Independently, a third set of programs was written to calculate the moments for the special queues, discussed in section 4. The moments were calculated using the methods given there. There results were used for testing the general programs for their correctness and numerical stability.

a. Generation of the Indices

The n-tuples $(j_1, j_2, \ldots, j_n)$ of indices are generated by a subroutine called INDEX, one at a time according to the algorithm discussed in [4]. A subroutine called PRNT then packs several sets of them into a single computer word for transfer to the external storage medium (tape or disk). Thus, the programs take into account the number of bits in each computer word and are therefore *machine dependent*. This machine dependence was necessary to make fully efficient use of the processing time and the memory storage.

An array of size 10,000 was reserved for the indices thus generated. The indices are formatted into records of as many complete sets of indices as possible. A complete set of indices for n is the set of all nonnegative n-tuples $(j_1, j_2, \ldots j_n)$ satisfying the condition

$$\sum_{i=1}^{n} i j_i = n.$$

Thus the first tape record contains indices for $n = 3, \ldots, 24$, the second for $n = 25$, 26, and 27, and

similarly for the further records. A magnetic tape was written containing the complete sets of indices for n from 3 to 49. This was the maximum number of complete sets of indices which would fit on a single 2,400-ft reel of 7-channel tape, recorded at 800 lines per inch.

b. Computation of Moments

The core of the moment computation routines is a subroutine called FORMY which evaluates the coefficients Y_{nr} in Fáa di Bruno's formula. The computations are organized so as to utilize the *packed* format of the indices without unpacking them. Furthermore, the subroutine is self contained i.e., does not call upon the system routines for functions, such as raising numbers to powers, etc. At each stage necessary indices are read into memory from the tape and the coefficients $Y_{nr}^{(1)}$, $Y_{nr}^{(2)}$, $Y_{nr}^{(3)}$, given by the Formulas (16), (19), and (22) for the model (a), are computed. After these computations, the memory space is released. This approach considerably shortened the running time and permitted efficient use of memory space, which becomes important after $n = 20$. Unfortunately, this approach necessitated writing the subroutine in assembly language.

The required input data for each moment calculation are N, the number of moments desired and λ the bunch arrival rate. In addition, pairs of subroutines, called ALPHA and ETA, which depend on the service time and bunch size distributions, are required to calculate the moments of these distributions.

The output first gives a summary of the input data, as well as the traffic intensity. If necessary the service time distribution is normalized to obtain $\alpha_1 = 1$. If such a normalization occurs, an appropriate message is given in the output. In addition, the moments of the service time and the bunch size distributions are given, along with the busy period moments.

c. Computational Accuracy and Running Times

In view of their great complexity and the large numbers of terms involved, the formulas involving several applications of Fáa di Bruno's formula were tested in several specific cases. The accuracy of the algorithm was ascertained by comparison with the numerical results, obtained by using the recurrence relations derived in section 4, in those cases where such tractable recurrences are available. Forty moments were calculated by both methods and an excellent agreement was found in all cases.

Other examples involved queues with gamma service time distributions and negative binomial bunch size distributions, among others. The parametric assumptions in these examples were made only to generate the higher moments of these distributions by the use of recursion formulas given in [3]. They neither limit nor simplify the application of the algorithm.

The results for these examples, as well as the computer program listings are given in a technical report, which is available upon request by writing to the authors at the Department of Statistics, Purdue University, W. Lafayette, Indiana 47907. Also available is a magnetic tape, containing the indices for Fáa di Bruno's formula for $3 \leq n \leq 49$. Use of this tape is, however, machine-dependent.

The running times required to generate the indices, as well as the moment calculation times are listed below. The time required by the system to load and set up the program was subtracted from the running time to give a better idea of the actual computation times.

Number of indices	Index generation (sec.)	Moment computation (sec.)
10	0.066	0.189
20	0.568	1.263
30	5.888	14.661
40	46.802	141.233
50	≈ 300	?

These are the central processing times for the CDC 6500 computer at Purdue University.

7. CONCLUSIONS

Our results adequately substantiate the claim that differentiation of the moment generating function using Fáa di Bruno's formula is entirely practical for obtaining the busy period moments of a single server queue with bunch arrivals. Up to 30 moments may be obtained using a very small amount of computer time; even up to 40 moments may be calculated with a moderate amount of computing time. Moments, beyond order 40 require considerable amounts of additional computing time.

The authors furthermore report that no problems of numerical instability or round off errors arose.

Our numerical investigations indicate that the moments of the busy period vary substantially with the bunch size and service time distributions, even for queues having the same traffic intensity. This shows that the preceding considerations are worthwhile, especially in view of the dearth of explicit and usable analytic results.

BIBLIOGRAPHY

[1] Dafermos, D., Stella, C., and Marcel F. Neuts, "A Single Server Queue in Discrete Time," Cahiers du Centre de Recherche Opérationnelle *13*, 23–40 (1971).

[2] Fáa di Bruno, C., *Théorie des Formes Binaires* (Librairie Brero, Succr de P. Marietti, Turin, Italy, 1876).

[3] Kendall, M. G. and A. Stuart, *The Advanced Theory of Statistics*, Vol. 1, 3rd ed. (Charles Griffin & Co., London, 1969).

[4] Klimko, E. M., "An Algorithm for Calculating Indices in Fáa di Bruno's Formula," forthcoming in B.I.T. (1973).

[5] Neuts, Marcel F., "The Single Server Queue in Discrete Time-Numerical Analysis I," Nav. Res. Log. Quart. *20.*, 297–304 (1973).

[6] Neuts, Marcel F. and David B. Wolfson, "Convexity of the Bounds Induced by Markov's Inequality," to appear, J. Stoch. Proc. and Appl. (1973).

[7] Takács, L., "The Transient Behavior of a Single Server Process with Poisson Input," *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 535–567. (Univ. of California Press, 1961).

[8] Takács, L., *Introduction to the Theory of Queues* (Oxford University Press, New York, 1962).

A QUEUE WITH WAITING TIME DEPENDENT SERVICE TIMES*

J. R. Callahan

Faculty of Administrative Studies
University of Manitoba
Canada

ABSTRACT

Sufficient conditions are developed for the ergodicity of a single server, first-come-first-serve queue with waiting time dependent service times.

INTRODUCTION

Queuing systems in which the service time of a customer is dependent on his waiting time in the queue have been largely ignored by queuing theorists. This type of dependence greatly increases the complexity of the state description of any model and its subsequent analysis.

On the other hand there are a number of real life queuing systems in which this kind of dependence is very important. During the production of steel, ingots are produced by pouring molten steel into molds. These ingots, after cooling sufficiently, are prepared for primary rolling in batteries of secondary heaters called soaking pits. At times there is no soaking pit space available and ingots must wait in a queue. The time necessary for the preparation of an ingot in a soaking pit is a function of its temperature when it is put in. While an ingot waits for space its temperature drops. Hence the system is an example of a queue with service times dependent on waiting times.

Another example is the queue of patients in the waiting list for beds in any hospital. As a patient waits, his condition and hence his subsequent treatment change.

Waiting time dependent service times are treated in an article by Sugawara and Takahashi [8]. Their analysis deals with the queue of steel which forms before the soaking pits during steel production.

They consider the service time of an ingot in the soaking pits as a deterministic, monotonically increasing function of its waiting time $f(w)$, with $f(0) > 0$ and $\lim_{w\to\infty} f(w) = f(\infty) < \infty$. Neglecting the interaction between the soaking pits and rolling mill, they consider the soaking pits as s independent service channels. They consider a renewal arrival process with expected interarrival time Eg_1 and first-come-first-serve pit queue discipline. They prove that if $s \cdot Eg_1 > f(\infty)$, then the queue is ergodic. If $Pr\{s_1 < f(0)\} = 0$, then no queue builds up at all even when $s \cdot Eg_1 < f(\infty)$.

Mudrov [5] published an article using a linear dependence between service time and waiting time in a single server first-come-first-serve queue, but circumvented the difficulties inherent in the model by considering "impatient" customers. If an impatient customer has to wait longer than a fixed time, he leaves the queue without getting served. In this way Mudrov found the average waiting time and hence the probability that a customer will be served.

Queues with state dependent service times, where state in this case means the number of customers in the system at the initiation of a service, are a related problem. Such systems have been studied

*This research was partially supported by a grant from the National Research Council of Canada.

by Conway and Maxwell [1], Rosenshine [7] and Harris [3, 4]. The state of the system and the waiting time of the customer entering service are related, but if the service time actually depends on waiting time, using a state dependent service time is not satisfactory. This is especially true when the customer interarrival times have a high variance.

The following theorem is presented as a contribution to the theory of queuing systems with waiting time dependent service times. It presents conditions for the ergodicity of a single server first-come-first-serve queue with a much more general dependency of service times on waiting times than that considered by Sugawara and Takahashi.

In the proof of the theorem, a theorem by Pakes[6] is used. His result is the following:

Consider $\{X_n\}$, an irreducible, aperiodic Markov chain with state space $\{0, 1, 2, \ldots\}$. Let $\delta_i = E\{X_{n+1} - X_n | X_n = i\}$. If $|\delta_i| < \infty$ for all i and if $\lim_{i\to\infty} \sup \delta_i < 0$, then $\{X_n\}$ is ergodic.

THEOREM: Consider a single server queue in which customers are served on a first-come-first-served basis. Let the time between successive arrivals at the queue be a discrete, nonnegative random variable, T. Consider that the service time, $S(i)$, of a customer is a discrete, nonnegative random variable which depends on i, the customers waiting time. (For simplicity, assume that $Pr\{S(i) - T = j\} > 0$, $j = -1, 0, 1$, for all i.)

Then if

$$\text{(i)} \quad E[S(i)] < \infty \qquad \text{for all } i = 0, 1, 2, \ldots$$

and

$$\text{(ii)} \quad \lim_{i\to\infty} \sup E[S(i)] < E[T],$$

the queue is ergodic.

PROOF: The waiting time of the nth customer to enter service, W_n, is given by the recursive relation,

$$W_n = \max\ (0, W_{n-1} + S(W_{n-1}) - T).$$

The sequence $\{W_n\}$ is a discrete parameter Markov chain defined on $\{0, 1, 2, \ldots\}$ which is irreducible and aperiodic because of the conditions on $Pr\{S(i) - T = j\}$, $\quad j = -1, 0, 1$.

From the recursive relation it can be seen that

$$E[W_n - W_{n-1} | W_{n-1} = i] \leqq E[S(i)] < \infty.$$

Since $\lim_{i\to\infty} \sup E[S(i)] < E[T]$ we have $\lim_{i\to\infty} \sup E[S(i) - T] < 0$ and hence there exists I_1 such that for all $i > I_1$,

$$E[S(i) - T] < -\phi \text{ for some } \phi > 0.$$

Now

$$E[W_n - W_{n-1} | W_{n-1} = i] = (-i) \cdot Pr\{S(i) - T \leqq -i\}$$

$$+ \sum_{k=-i+1}^{\infty} k \cdot Pr\{S(i) - T = k\} \leqq \sum_{k=-i+1}^{\infty} k \cdot Pr\{S(i) - T = k\}$$

so that

$$E[S(i) - T] \leqq E[W_n - W_{n-1} | W_{n-1} = i]$$

$$\leqq \sum_{k=-i+1}^{\infty} k \cdot Pr\{S(i) - T = k\}.$$

Since $E[T] < \infty$, we have

$$\lim_{i \to \infty} (-i) Pr \cdot \{S(i) - T \leqq -i\} = 0$$

and hence as $i \to \infty$

$$\sum_{k=-i+1}^{\infty} k \cdot Pr\{S(i) - T = k\} \to E[S(i) - T].$$

That is, there exists I_2 such that for all $i > I_2$

$$E[S(i) - T] + \phi/2 > E[W_n - W_{n-1} | W_{n-1} = i].$$

Thus for $i > I = \max(I_1, I_2)$

$$E[W_n - W_{n-1} | W_{n-1} = i] < -\phi/2.$$

That is

$$\lim_{i \to \infty} \sup E[W_n - W_{n-1} | W_{n-1} = i] > 0.$$

Thus $\{W_n\}$ fulfills all of the conditions of Pakes' theorem, and hence is ergodic. Q.E.D.

The condition placed on the interarrival and service times, that they be discrete, is a serious one in theory, but the kind of continuous distributions used in practice can be adequately represented with discrete time approximations. Dafermos and Neuts [2], in a recent article, mention the usefullness of studying hitherto intractable queuing problems using discrete time models.

Moreover, it should be true that if the interarrival and service times have general distributions, the queue is ergodic so long as $E[S(w)] < \infty$ and $\lim_{w \to \infty} \sup E[S(w)] < E[T]$, where w is a continuous variable.

CONCLUSION

It has been determined for a single server first-come-first-serve queue with discrete arrival and service processes in which service times depend on waiting times, that if there exists some waiting time w_0 such that when a customer waits for w_0 or longer, its average service time is smaller than the average interarrival time of customers, the queue is ergodic. It is hoped that this analysis will promote research in a long ignored, but important class of queuing problems.

ACKNOWLEDGMENTS

The author would like to acknowledge the financial support of the National Research Council of Canada during his research.

REFERENCES

[1] Conway R. W. and W. L. Maxwell, "A Queuing Model with State Dependent Service Rates," J. Indust. Eng. *12*. 132–136 (1962).

[2] Dafermos S. C. and M. F. Neuts, "A Single Server Queue in Discrete Time," Cahiers du Centre de Recherche Opérationelle *13*, 23–40 (1971).

[3] Harris C. M., "Queues with State Dependent Service Rates," Operations Research *15*, 117–130 (1967).

[4] Harris, C. M., "Some Results for Bulk-Arrival Queues with State Dependent Service Times," Man. Sci: Theory *16*, 313–326 (1970).

[5] Mudrov V. I., "Queuing with Impatient Customers and Variable Service Time, Linearly Dependent on the Queuing Time of the Customer," Problemy Kobernetiki *5*, 283–385 (1961).

[6] Pakes A. G., "Some Conditions for Ergodicity and Recurrence of Markov Chains," Operations Research *17*, 1058–1061 (1969).

[7] Rosenshine M., "Queues with State Dependent Service Times," Transpn. Res. *1*, 97–104, (1967).

[8] Sugawara S. and M. Takahashi, "On Some Queues Occurring in an Integrated Iron and Steel Works," J of O.R. Soc. of Japan *8*, 16–23 (Sept. 1965).

AN ALGORITHM FOR SEPARABLE PIECEWISE CONVEX PROGRAMMING PROBLEMS

Richard M. Soland

The University of Texas at Austin

ABSTRACT

We present a branch and bound algorithm to solve mathematical programming problems of the form: Find $x=(x_1, \ldots, x_n)$ to minimize $\Sigma\varphi_{i0}(x_i)$ subject to $x\epsilon G$, $l \leq x \leq L$, and $\Sigma\varphi_{ij}(x_i) \leq 0, j=1, \ldots, m$. With $l=(l_1, \ldots, l_n)$ and $L=(L_1, \ldots, L_n)$, each φ_{ij} is assumed to be lower semicontinuous and piecewise convex on the finite interval $[l_i, L_i]$. G is assumed to be a closed convex set. The algorithm solves a finite sequence of convex programming problems; these correspond to successive partitions of the set $C=\{x|l \leq x \leq L\}$ on the basis of the piecewise convexity of the problem functions φ_{ij}. Computational considerations are discussed, and an illustrative example is presented.

I. INTRODUCTION

We consider the following problem: Find a vector $x=(x_1, \ldots, x_n)$ which minimizes

$$\varphi_0(x) \equiv \sum_{i=1}^{n} \varphi_{i0}(x_i)$$

subject to

(Problem P)
$$x\epsilon G,$$
$$x\epsilon C \equiv \{l \leq x \leq L\},$$
$$\varphi_j(x) \equiv \sum_{i=1}^{n} \varphi_{ij}(x_i) \leq 0, \qquad j=1, \ldots, m.$$

It is assumed that the set G is closed and convex, that $-\infty < l_i < L_i < +\infty$, and that each φ_{ij} is lower semicontinuous on the interval $[l_i, L_i]$. Let $M=\{1, 2, \ldots, m\}$, let $M^*=\{0, 1, \ldots, m\}$, let

$$H=\{x|x\epsilon C;\ \varphi_j(x) \leq 0,\ j\epsilon M\},$$

and assume that $G \cap H$ is non-empty. These assumptions are sufficient to ensure that φ_0 attains its minimum over the set $G \cap H$. It is further assumed that each φ_{ij} is a piecewise convex function on $[l_i, L_i]$. More precisely, we suppose that there are a finite number of points $e_{ijp}(p=0, 1, \ldots, r_{ij})$ with the following properties:

(a) $l_i=e_{ij0} \leq e_{ij1} \leq e_{ij2} \leq \ldots \leq e_{ijr_{ij}}=L_i$,

(b) if $e_{ijp} < e_{ijp+1}$, then φ_{ij} is convex (and therefore continuous) on the open interval (e_{ijp}, e_{ijp+1})

The result of these assumptions is that the interval $[l_i, L_i]$ can be partitioned into a finite number of subintervals, over each of which φ_{ij} is convex. This type of partition of the n intervals $[l_i, L_i]$, after adjustment for the various values of j, forms the basis of a branch and bound algorithm for the solution of Problem P.

We shall attempt to indicate, very briefly, how such piecewise convex objective and constraint functions may arise. Figure 1 shows a possible production cost curve that is piecewise convex. After a fixed cost at $x_i=0$, the marginal production cost is nonincreasing until $x_i=300$, whereupon it begins to increase rapidly. In this case we may take $r_{i0}=3$, with $e_{i00}=0$, $e_{i01}=100$, $e_{i02}=200$, $e_{i03}=500$. As a second example, suppose the constraint $\varphi_1(x) \leq 0$ is equivalent to an upper limit on the number of hours of labor that may be utilized in the production of several products. Suppose the method of production of product i depends upon the quantity that will be produced. Specifically, suppose 1 hour of labor per unit is required if fewer than 200 units are produced. If the number of units produced is at least 200 but less than 1,000, a semi-automated process requiring 0.5 hour of labor per unit will be used. If the number produced lies between 1,000 and 3,000, an automated process requiring 0.25 hour of labor per unit will be used. The resulting piecewise linear curve of labor hours $\varphi_{i1}(x_1)$ versus units produced x_1 is shown in Figure 2.

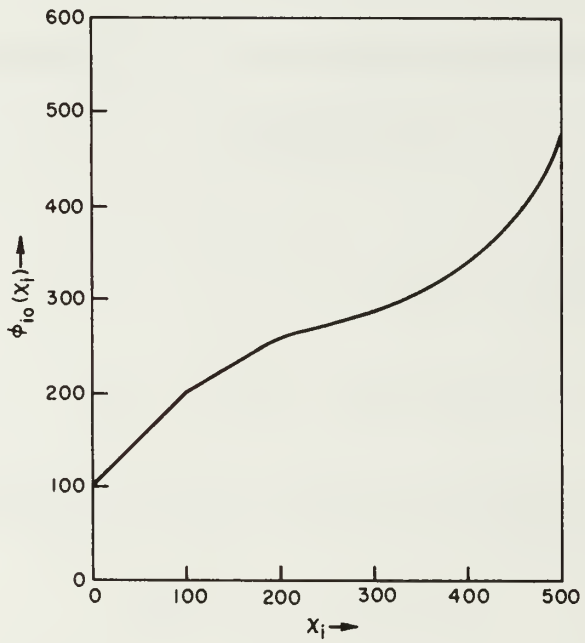


FIGURE 1. A piecewise convex production cost function.

The algorithm we propose for the solution of Problem P is an extension and generalization of one given by Jones and Soland [8] for the case in which the restrictions $\varphi_j(x) \leq 0$, $j \in M$, are not present. The idea of treating programming problems containing piecewise convex functions by branch and bound was first proposed by Beale [3] for a transportation problem with decreasing marginal costs. Mitra [9, 10] treated the fixed charge problem and the fixed charge transportation problem by methods very similar to the algorithm of Jones and Soland.

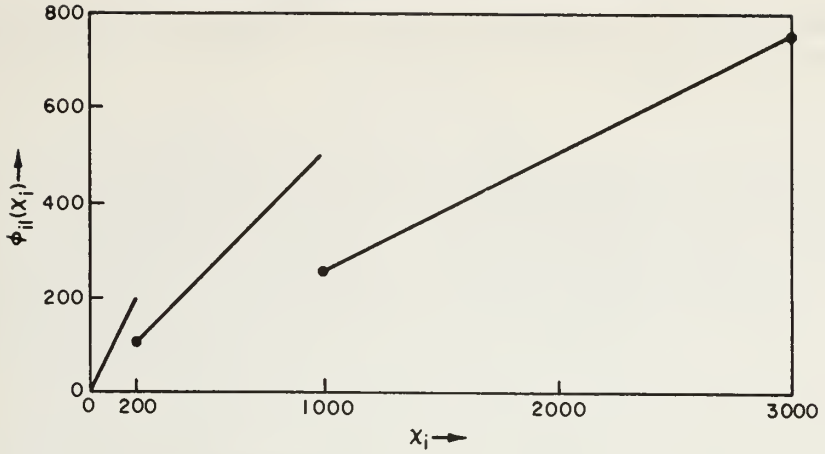


FIGURE 2. A piecewise convex function of labor hours vs units produced.

The statement of Problem P is almost identical to that of the problem considered by Soland [12]. Our problem here is less general, however, in that the algorithm of [12] applies to problems in which the φ_{ij} are not piecewise convex. The present algorithm, however, converges to an optimal solution of Problem P after solving a *finite* number of convex programming problems, whereas the algorithm of [12] produces, in general, an *infinite* sequence of convex problems, the solutions to which converge to an optional solution of the original problem. In addition, the algorithm of [12] requires a potentially inefficient branching (or partition) rule in order to guarantee convergence when the problem functions have discontinuities.

Our use of the term *separable programming* differs somewhat from the common usage in that we allow the functions defining G to be nonseparable. More importantly, the usual approach to the solution of separable programming problems uses piecewise linear functions to approximate the given objective and constraint functions. In the usual λ-formulation these separable programming problems can then be represented as mixed-integer linear programming problems (see Beale [4]). The algorithm of this paper can also be used to solve such piecewise linear approximating problems.

A brief outline of the paper is as follows. In section 2 we do some preliminary analysis of the problem, and in section 3 we present the algorithm. Section 4 contains a discussion of some computational considerations and extensions and section 5 treats an illustrative example.

II. PRELIMINARIES

We previously stated that the interval $[l_i, L_i]$ can be partitioned into a finite number of subintervals, over each of which φ_{ij} is convex. The object of this section is to obtain a partition of $[l_i, L_i]$ into a finite number of subintervals, over each of which *all* φ_{ij}, $j \epsilon M^*$, are convex.

For each i and for each $j \epsilon M^*$, define a finite set A_{ij} as follows: (1) A_{ij} contains l_i^- and L_i^+, e.g., if $l_i = 0$ and $L_i = 5$, then A_{ij} contains the elements 0^- and 5^+. (2) If φ_{ij} is continuous at some $e_{ijp} \epsilon (l_i, L_i)$, then A_{ij} contains the element e_{ijp}. (3) If φ_{ij} is *not* left (right) continuous at some $e_{ijp} \epsilon [l_i, L_i]$ then A_{ij} contains the element e_{ijp}^- (e_{ijp}^+).

An example should aid in making this definition clear. Figure 3 shows φ_{ij}, $j \in M^*$, for $m=2$. We may choose the e_{ijp} as follows:

	e_{ij0}	e_{ij1}	e_{ij2}	e_{ij3}	e_{ij4}
$j=0$	0	1	2	4	
$j=1$	0	1	2	3	4
$j=2$	0	1	3	4	

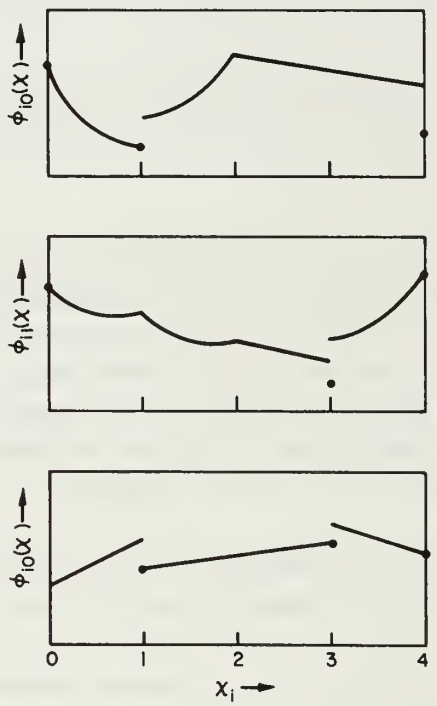


FIGURE 3. Three piecewise convex functions of x_i.

(We *could*, of course, include additional points, e.g., take $e_{i03}=3.5$, $e_{i04}=4$, but there is nothing to be gained, and something to be lost, by including such unnecessary points. Each r_{ij} should be as small as possible.)

For this example the reader may then verify the following:

$$A_{i0}=\{0^-, 1^+, 2, 4^-, 4^+\},$$

$$A_{i1}=\{0^-, 1, 2, 3^-, 3^+, 4^+\},$$

$$A_{i2}=\{0^-, 0^+, 1^-, 3^+, 4^+\}.$$

Let

$$A_{i*} = \bigcup_{j=1}^{m} A_{ij}.$$

Define the ordered set A_i as follows: (1) If $a^- \epsilon A_{i*} (a^+ \epsilon A_{i*})$, then $a^- \epsilon A_i (a^+ \epsilon A_i)$. (2) If $a \epsilon A_{i*}$, but $a^- \notin A_{i*}$, then $a_i^+ \epsilon A_i$. A_i is ordered in the natural arithmetic order as follows: (3) The first (last) element of A_i is $l_i^- (L_i^+)$. (4) If $b^* \epsilon A_i$ and $c^* \epsilon A_i$, where "*" is either "−" or "+," and $b < c$, then b_i^* preceeds c_i^* in the set A_i. (4) If $a^- \epsilon A_i$ and $a^+ \epsilon A_i$, then a^- is the immediate predecessor of a^+.

For the example of Figure 3 we have

$$A_{i*} = \{0^-, 0^+, 1^-, 1, 1^+, 2, 3^-, 3^+, 4^-, 4^+\},$$

$$A_i = \{0^-, 0^+, 1^-, 1^+, 2^+, 3^-, 3^+, 4^-, 4^+\}.$$

Let the number of elements of A_i be $r_i + 1$, and label the (ordered) elements of A_i as $a_{i0}, a_{i1}, a_{i2}, \ldots, a_{ir_i}$. Define the interval I_{ip}, $p = 1, \ldots, r_i$ as follows: The left and right endpoints of I_{ip} are the numerical values (ignoring the −'s and +'s) of $a_{i,p-1}$ and a_{ip} respectively. If $a_{i,p-1}$ has a "−" sign ("+" sign) the lefthand endpoint is included (excluded). If a_{ip} has a "−" sign ("+" sign) the righthand endpoint is excluded (included). For the example of Figure 3 we have $r_i = 8$ and

$$\begin{aligned} I_{i1} &= [0, 0], & I_{i2} &= (0, 1), \\ I_{i3} &= [1, 1], & I_{i4} &= (1, 2], \\ I_{i5} &= (2, 3), & I_{i6} &= [3, 3], \\ I_{i7} &= (3, 4), & I_{i8} &= [4, 4]. \end{aligned}$$

It is evident from the construction of the intervals I_{ip} that they form a disjoint union of $[l_i, L_i]$, i.e.,

(1) $$\bigcup_{p=1}^{r_i} I_{ip} = [l_i, L_i],$$

(2) $$I_{ip} \cup I_{iq} = \emptyset, \quad p \neq q.$$

Equally important is the fact that if x_i is confined to any I_{ip}, *all* $\varphi_{ij}(x_i)$, $j \epsilon M^*$, are continuous convex functions of x_i. Suppose, now, we restrict each x_i to a particular I_{ip_i} and form Problem $\hat{\mathrm{P}}$: Find an $x = (x_1, \ldots, x_n)$ which minimizes

$$\sum_{i=1}^{n} \varphi_{i0}(x_i),$$

subject to

$$x \epsilon G,$$

(Problem $\hat{\mathrm{P}}$) $$\sum_{i=1}^{n} \varphi_{ij}(x_i) \leq 0, \qquad j = 1, \ldots, m,$$

$$x \epsilon I_{ip_i}, \qquad i = 1, \ldots, n.$$

The constraint set of Problem $\hat{P}$ is convex, and on this constraint set the objective function is continuous and convex. Thus any local solution of Problem $\hat{P}$ is also a global solution, and readily available computational methods (such as SUMT [7]) could be used to solve Problem $\hat{P}$. It is quite possible, of course, that Problem $\hat{P}$ has no solution because either (1) there is no feasible solution, or (2) the infimum of the objective function occurs at a boundary point that is not feasible for Problem $\hat{P}$. These possible difficulties are circumvented in the algorithm to be described later.

We may suppose that it is possible to find an optimal solution of Problem $\hat{P}$ or establish that none exists. The number of unique problems of which Problem $\hat{P}$ is just one example is clearly

$$R \equiv \prod_{i=1}^{n} r_i \tag{3}$$

since each p_i may be chosen as any one of r_i values. Conceptually, then, Problem P may be solved by setting up and solving each of the R problems of the form Problem $\hat{P}$ and then choosing the best of the solutions found.

The branch and bound algorithm of section 3 implicitly considers all of these R problems, but is designed with the aim of solving only a relatively small number of convex programming problems.

In order to avoid any difficulties due to a constraint set that is not closed we make the following definitions:

(4) I_{ip}^{*} is the closure of I_{ip},

(5) φ_{ijp} is the continuous extension of φ_{ij} to the closed interval I_{ip}^{*}.

With these definitions we could replace Problem $\hat{P}$ with the following: Find an x which minimizes

$$\sum_{i=1}^{n} \varphi_{iop_i}(x_i),$$

subject to

$$x \in G$$

(Problem $\hat{P}^*$)
$$\sum_{i=1}^{n} \varphi_{ijp_i}(x_i) \leq 0, \qquad j = 1, \ldots, m,$$

$$x \in I_{ip_i}^{*}, \qquad i = 1, \ldots, n.$$

Problem $\hat{P}^*$ involves the minimization of a continuous convex function over a closed convex set and therefore has an optimal solution or else is infeasible. An optimal solution to Problem $\hat{P}^*$ may not be feasible for Problem $\hat{P}$, but it *is* feasible for Problem P because each φ_{ij} is lower semicontinuous. For the same reason, the actual objective function value at a point that solves Problem $\hat{P}^*$ is less than or equal to the optimal value of Problem $\hat{P}^*$. If one wished to solve Problem P by *explicitly* considering all R problems, he would solve R problems of the form Problem $\hat{P}^*$ and choose the solution with smallest objective function value.

III. THE BRANCH AND BOUND ALGORITHM

In this section we present a branch and bound algorithm (see [1, 2, 11]) for the solution of Problem P. The algorithm considers subsets of the feasible solutions of Problem P, finds lower bounds on the optimal solution values in such subsets by solving convex programming problems, and eventually identifies an optimal solution. It involves a finite sequence of convex programming problems.

We will phrase the algorithm in terms of a branch and bound tree, as is frequently done with branch and bound algorithms. Let the nodes of the tree be $N^0, N^1, \ldots$, with N^0 as the initial node. The number of nodes created at any stage of the algorithm is equal to r_i for some i. Each node N^k represents a subset of solutions T^k of Problem P that is characterized by an n-dimensional vector σ^k. The ith entry of σ^k has the possible values $0, 1, 2, \ldots, r_i$. An entry of 0 means that T^k consists of solutions in which x_i is merely restricted to $[l_i, L_i]$; an entry of $p > 0$ means that T^* consists of solutions in which x_i is restricted to I_{ip}. Thus T^0, the set of all solutions of Problem P, is characterized by a vector of zeros. $\sigma^k = (0, 3, 1, 0)$ represents the subset T^k in which $x_1 \epsilon [l_1, L_1]$, $x_2 \epsilon I_{23}$, $x_3 \epsilon I_{31}$, and $x_4 \epsilon [l_4, L_4]$.

Let $LB(N^k)$ be a lower bound on the optimal value of φ_0 over T^k; it is equal to the optimal value of the following convex programming problem obtained from Problem P. For each i such that σ_i^k, the ith entry of σ^k, is $p_i > 0$, the variable x_i is restricted to $I^*_{ip_i}$ and each φ_{ij}, $j \epsilon M^*$, is replaced by φ_{ijp_i}. For each i such that $\sigma_i^k = 0$, x_i is restricted to $[l_i, L_i]$ and each φ_{ij}, $j \epsilon M^*$, is replaced by an underestimating function ψ_{ij} that has the following properties:

$$\psi_{ij}(x_i) \leqslant \varphi_{ij}(x_i) \text{ for } x_i \epsilon [l_i, L_i], \tag{6}$$

$$\psi_{ij} \text{ is a continuous convex function on } [l_i, L_i]. \tag{7}$$

(Possible choices for the ψ_{ij} will be discussed in section 4.) The optimal value of the resulting convex programming problem is defined to be $LB(N^k)$ (we set $LB(N^k) = +\infty$ if the problem is infeasible). The validity of this lower bound follows from the definition of the ψ_{ij} and the fact that $I_{ip_i} \subset I^*_{ip_i}$.

An intermediate node at any stage of the branch and bound tree is one that has not yet been branched. Branching from an intermediate node N^k is carried out as follows: Choose any i such that $\sigma_i^k = 0$, and branch node N^k to r_i new nodes, the pth of which corresponds to a vector σ that is identical to σ^k except that $\sigma_i = p$. That is, in the pth of the r_i subsets into which T^k is partitioned, x_i is restricted to I_{ip}. We offer, as one possible way of choosing i, the following branching rule:

Branching Rule. Computation of $LB(N^k)$ yields a solution vector x^k. Choose any i that maximizes the difference

$$\varphi_{ij}(x_i^k) - \psi_{ij}(x_i^k).$$

Here i takes the values $1, \ldots, n$ such that $\sigma_i^k = 0$, but j is restricted to the value 0 and the integers $j \epsilon M$, such that $\varphi_j(x^k) > 0$. That is, only the objective function and the violated constraints are considered in choosing i. (This branching rule is patterned after the weak refining rule of [12].)

The algorithm produces a sequence of (not necessarily feasible) points $\{x^k : k = 0, 1, 2, \ldots\}$, where x^k is found at node N^k (if the convex programming problem at node N^k is infeasible we may take $x^k = x^{k-1}$). Let

$$UB_\varphi^k = \min \{\varphi_0(x^h) : h = 0, 1, \ldots, k; x^k \epsilon H\},$$

and let $UB_x^k = x^h$, where $UB_\varphi^k = \varphi_0(x^h)$ (if none of the vectors $x^h, h = 0, \ldots, k$, are feasible for Problem P we may set UB_x^k arbitrarily and $UB_\varphi^k = \varphi_0(UB_x^k)$ if $UB_x^k \in G \cap H$ and $UB_\varphi^k = +\infty$ otherwise). UB_φ^k is an upper bound on the optimal solution value of Problem P and UB_x^k is the best solution found through node N^k.

Now we may state the algorithm in Steps A, B, and C.

Step A. Find $LB(N^0)$, UB_φ^0, UB_x^0. Go to Step B.

Step B. Suppose N^q is the highest numbered intermediate node. Find k^* such that $LB(N^{k*})$ is the minimum of $LB(N^h)$ over all intermediate nodes N^h. If $LB(N^{k*}) \geq UB_\varphi^q$, UB_x^q is an optimal solution. Otherwise, set $s = k^*$ and go to Step C.

Step C. Let i_s be the index of the variable on which to branch node N^s. Let $t = r_{is}$. Branch node N^s to nodes $N^{q+1}, \ldots, N^{q+t}$. Find $LB(N^{q+1}), \ldots, LB(N^{q+t})$. Find UB_x^{q+t} and UB_φ^{q+t}. Go to Step B.

It should be clear that if the algorithm does terminate at Step B, then UB_x^q is an optimal solution. This follows directly from the validity of the lower and upper bounds and the fact that no feasible solution of Problem P is ever excluded from consideration. The algorithm must be finite since by the nth level of the tree the lower bound determined for an intermediate node will equal or exceed the solution value of the vector found at that node.

IV. COMPUTATIONAL CONSIDERATIONS AND EXTENSIONS

In this section we briefly comment on a number of factors affecting the computational usefulness of the algorithm; some of these comments pertain to a wide class of branch and bound algorithms. Two simple extensions of the algorithm are presented also.

A key question concerns the ψ_{ij} underestimating functions. What functions should be used? Given the requirements (6) and (7) that ψ_{ij} be continuous and convex on $[l_i, L_i]$ and $\psi_{ij}(x_i) \leq \varphi_{ij}(x_i)$ for $x_i \in [l_i, L_i]$, it is clearly desirable that $\psi_{ij}(x_i)$ be as large as possible, i.e., that $\psi_{ij}(x_i)$ be as close to $\varphi_{ij}(x_i)$ as possible. The lower bounds $LB(N^k)$ will then be as tight as possible, given the current method of calculating the bounds, and it is then possible that a greater number of intermediate nodes need never be branched. Requiring ψ_{ij} to be the greatest convex function that underestimates φ_{ij} means that ψ_{ij} is the *convex envelope* of φ_{ij} over $[l_i, L_i]$. (More precise definitions of the convex envelope of a function, and their equivalence, are given by Falk [6].) It is sometimes difficult to calculate the convex envelope of a given function, however. Also, the convex envelope may not have a derivative everywhere in $[l_i, L_i]$, and this may be computationally undesirable. Jones and Soland [8] used as an underestimate the greatest linear function agreeing with φ_{ij} at the lower limit l_i. It is worthwhile repeating that any underestimate possessing properties (6) and (7) will do, and that each ψ_{ij} need only be chosen once.

Most computational methods of convex nonlinear programming require, or at least make it desirable, that the problem functions have first derivatives or continuous first derivatives. We might, therefore, require that φ_{ij} also have a first derivative or continuous first derivative on the open interval (e_{ijp}, e_{ijp+1}). We would then put the same requirement on the underestimating functions ψ_{ij} over the interval $[l_i, L_i]$.

As in most branch and bound algorithms, this algorithm may be stopped when a solution has been found that is sufficiently close to an optimal one in objective function value. Comparison should be made with $UB_\varphi^q - \epsilon$ instead of UB_φ^q in Step B of the algorithm, where ϵ is a positive tolerence. If UB_φ^q is positive, ϵ may be calculated as $[UB_\varphi^q][\rho(1+\rho)^{-1}]$, where ρ is the fractional tolerence to within which an optimal solution must be found. Use of a positive ϵ, as opposed to zero, is desirable to avoid difficulties due to computer roundoff error.

In branching a node N^k, r_i new nodes are created, and a lower bound must be found for each. It may be computationally efficient to begin each associated convex programming problem from x^k, the solution vector for the problem yielding $LB(N^k)$. This problem and its r_i descendents do not differ very much, so starting at x^k may lead to faster solutions. The point x^k itself will be infeasible for at least $r_i - 2$ of the descendent problems, however.

If a dual feasible method is used to find the lower bound $LB(N^k)$, a lower bound on $LB(N^k)$ itself is available during the minimization. If this ever equals or exceeds UB_φ^{k-1}, then $LB(N^k)$ certainly will, and so the minimization may be stopped.

Computationally, this algorithm would seem to hold some promise if the r_i's are reasonably small. If they are quite large, then it is probable that some special structure is present in the problem and some method of solution other than the present algorithm might be more practical.

Separable equality constraints of the form

$$\varphi_j(x) = \sum_{i=1}^{n} \varphi_{ij}(x_i) = 0$$

can easily be handled by the algorithm. Merely replace this constraint by the two inequality constraints

$$\varphi_j(x) \leqslant 0, \qquad -\varphi_j(x) \leqslant 0.$$

Each φ_{ij} must be continuous and piecewise linear in order that both φ_{ij} and $-\varphi_{ij}$ be lower semicontinuous and piecewise convex.

We may add a nonseparable (or separable) continuous convex function $\theta_j(x)$ to each function $\varphi_j(x)$ of Problem P and use the same algorithm. The only changes required are (1) at node N^k each $\theta_j(x)$ must be added appropriately to form the convex programming problem whose optimal solution is $LB(N^k)$, and (2) $\theta_0(x^h)$ must be added to $\varphi_0(x^h)$ in computing UB_φ^k.

V. AN ILLUSTRATIVE EXAMPLE

This example is patterned after the one given on pages 85–86 of [5]. A firm processes a specific raw material using two types of equipment, called stills and retorts. Four different products may be made from the raw material, and the firm wishes to determine how many units of each product to make so as to maximize its profit, or minimize the negative of its profit. There are three constraints; (1) only 1,500 tons of raw material are available, (2) the firm's still capacity must not be exceeded, and (3) the firm's retort capacity must not be exceeded.

Let x_i, $i = 1, 2, 3, 4$, be the number of units of product i to be made. We assume that one unit of any of the four products requires 100 tons of raw material. The first constraint, the one that defines G, is thus

$$G\colon 100x_1 + 100x_2 + 100x_3 + 100x_4 \leqslant 1{,}500.$$

The second and third constraints, those on still and retort capacity, involve piecewise linear (and therefore piecewise convex) functions of the x_i. We will explain how these arise by treating the case of product 1. Three somewhat different production methods are available for product 1. The first will be used if $x_1 < 5$, the second will be used if $5 \leqslant x_1 < 10$, and the third will be used if $10 \leqslant x_1 \leqslant 15$. The percentage of still capacity used to produce x_1 units of product 1 is $8x_1$ if the first production method

is used ($x_1 < 5$), $7x_1$ if the second is used ($5 \leq x_1 < 10$), and $6x_1$ if the third is used ($10 \leq x_1 \leq 15$). Define $\varphi_{11}(x_1)$ as the percentage of still capacity used to produce x_1 units of product 1. Then

$$\varphi_{11}(x_1) = \begin{cases} 8x_1 & \text{if } 0 \leq x_1 < 5, \\ 7x_1 & \text{if } 5 \leq x_1 < 10, \\ 6x_1 & \text{if } 10 \leq x_1 \leq 15. \end{cases}$$

With $\varphi_{i1}(x_i)$, $1 = 2, 3, 4$, similarly defined as the percentage of still capacity used to produce x_i units of product i, the constraint we want is

$$\sum_{i=1}^{4} \varphi_{i1}(x_i) \leq 100.$$

This is not quite in the required form because of the nonzero right-hand side, but this can be easily remedied by subtracting 100 from one of the φ_{i1}; we will ignore this minor chore.

Define $\varphi_{i2}(x_i)$, $i = 1, \ldots, 4$, as the percentage of retort capacity used to produce x_i units of product i. $\varphi_{12}(x_1)$ is a piecewise linear function of x_1 for the same reasons that $\varphi_{11}(x_1)$ is. The constraint on retort capacity is then

$$\sum_{i=1}^{4} \varphi_{i2}(x_i) \leq 100.$$

Let $\varphi_{i0}(x_i)$, $i = 1, \ldots, 4$, be the negative of the net profit obtained by producing and selling x_i units of product i. For the reasons cited above the φ_{i0} functions are also assumed piecewise linear. Some additional changes in slope and value of the φ_{i0} functions are due to economies of scale and setup costs, respectively. All of the φ_{ij} functions for this problem are shown in Figure 4.

We also need upper limits on the x_i (the lower limits are clearly zero). From the inequality defining G we have $x_i \leq 15$, and from the fact that $\varphi_{42}(x_4) > 100$ for $x_4 > 10$ we may set $L_4 = 10$. The overall problem statement is then: Find $x_1, \ldots, x_4$ to minimize

$$\sum_{i=1}^{4} \varphi_{i0}(x_i),$$

subject to

$$G: \sum_{i=1}^{4} 100x_i \leq 1{,}500,$$

$$\sum_{i=1}^{4} \varphi_{ij}(x_i) \leq 100, \qquad j = 1, 2,$$

$$0 \leq x_i \leq 15, \qquad i = 1, 2, 3,$$

$$0 \leq x_4 \leq 10.$$

In solving this example by the algorithm, the ψ_{ij} underestimating functions were chosen to be the linear underestimates with greatest slope that agree with the φ_{ij} at the lower limits l_i. They are shown in Figure 4. Each convex programming problem solved was therefore a linear programming problem. The branching rule presented along with the algorithm was used.

From Figure 4 the A_i sets are seen to be

$$A_1 = \{0^-, 0^+, 5^-, 10^-, 15^+\},$$
$$A_2 = \{0^-, 0^+, 3^+, 6^-, 9^+, 15^+\},$$
$$A_3 = \{0^-, 0^+, 4^+, 8^-, 11^+, 15^+\},$$
$$A_4 = \{0^-, 0^+, 3^+, 5^-, 10^+\}.$$

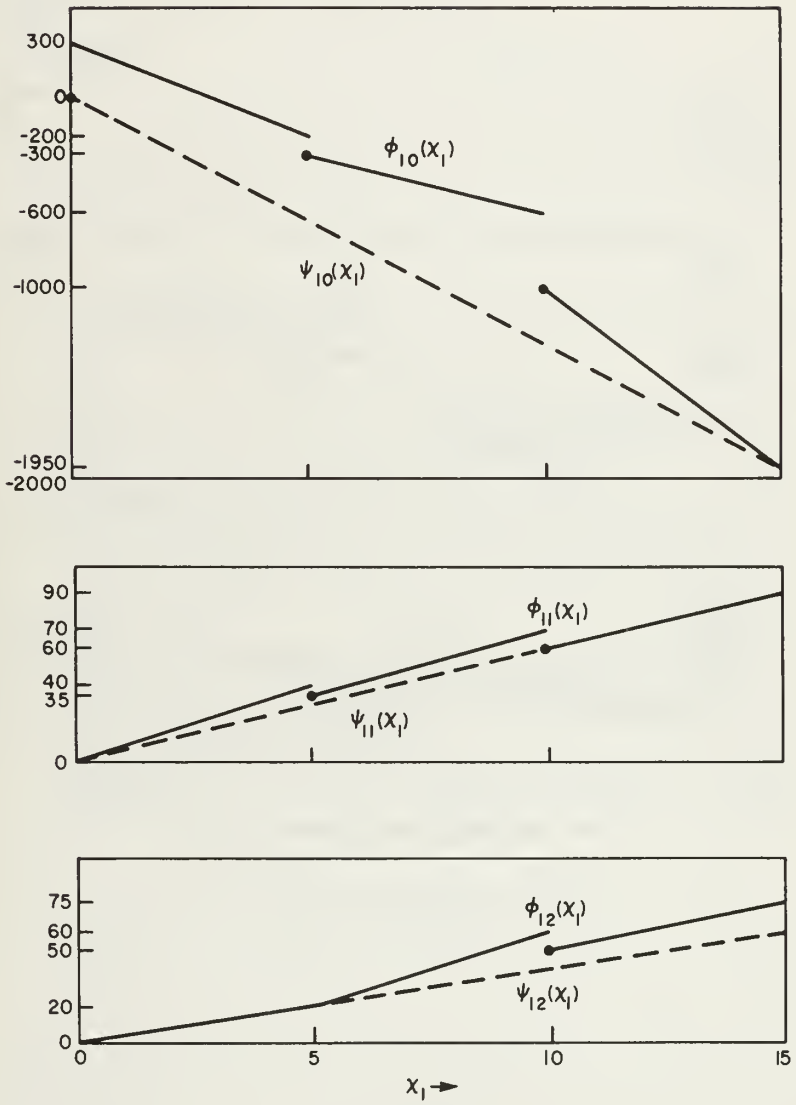
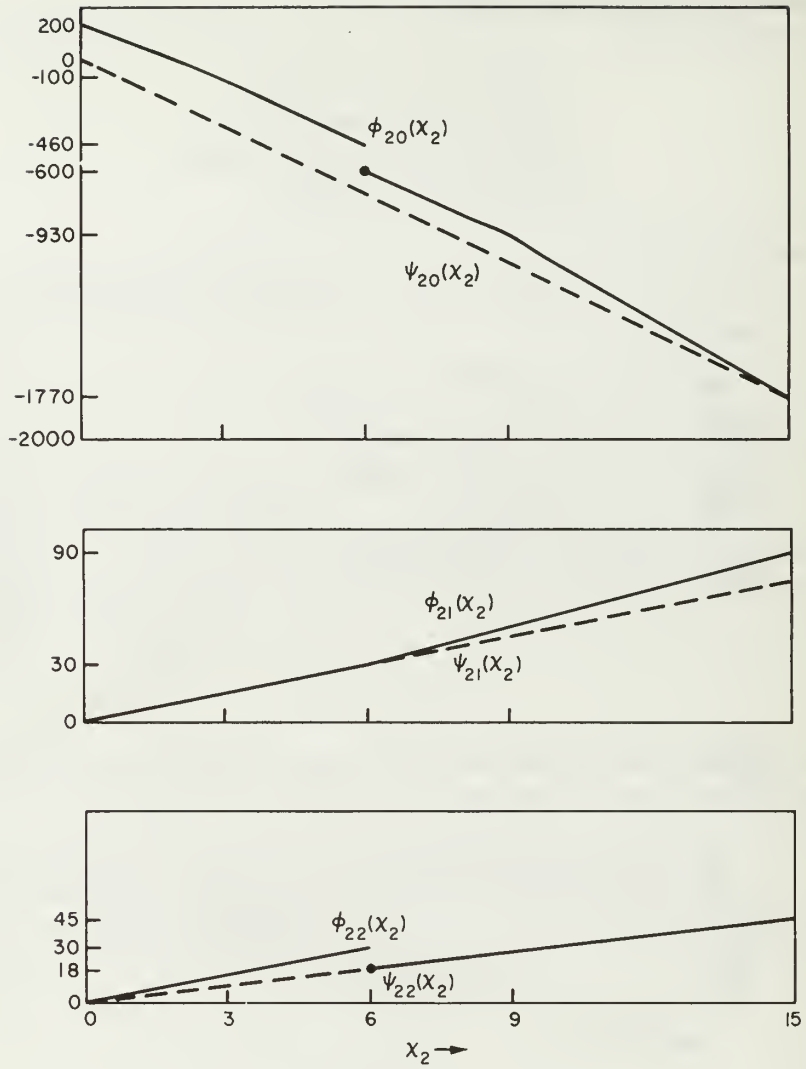
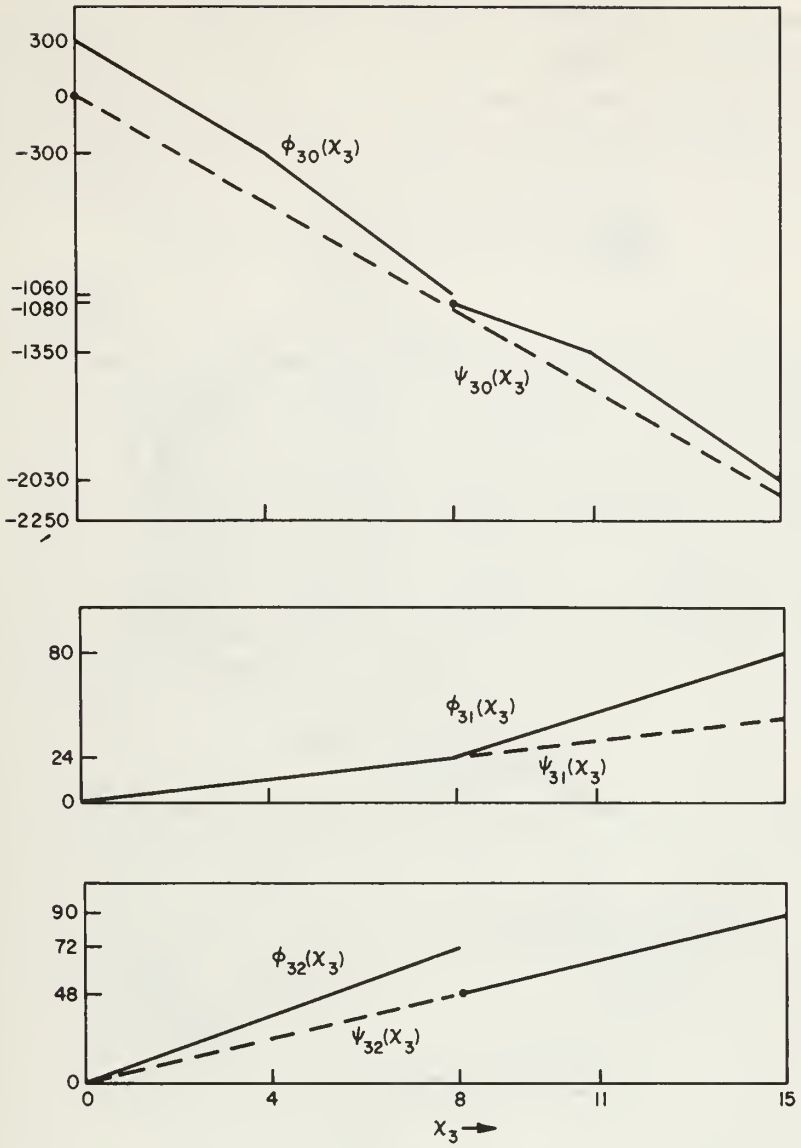


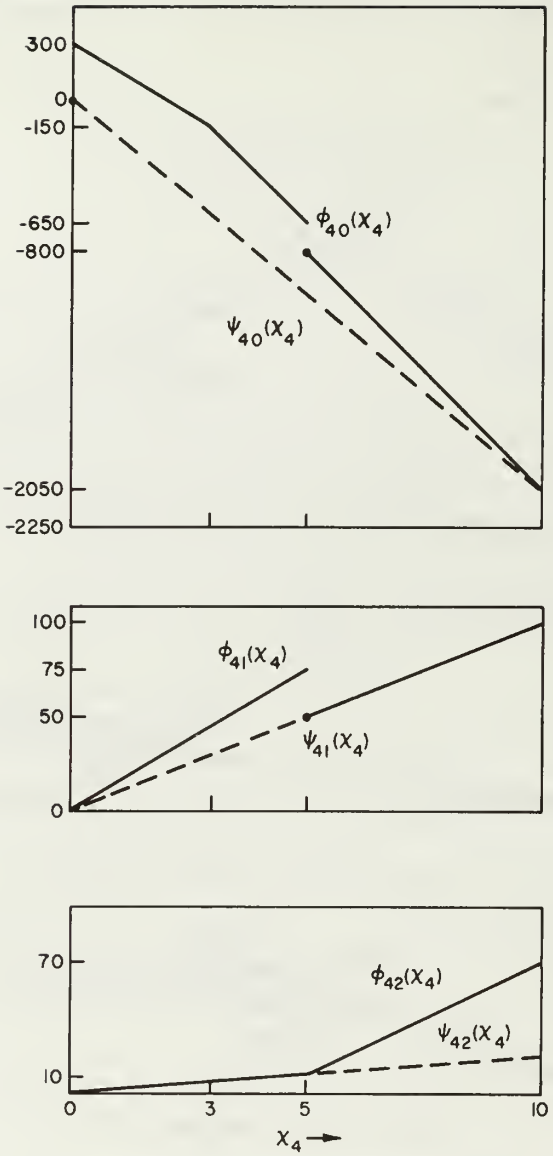
FIGURE 4 (a). ϕ_{1j} and ψ_{1j} for the illustrative example.



(b). ϕ_{2j} and ψ_{2j} for the illustrative example.



(c). ϕ_{3j} and ψ_{3j} for the illustrative example.



(d). ϕ_{4j} and ψ_{4j} for the illustrative example.

We therefore have $R = (4)(5)(4) = 400$. The algorithm solved 10 linear programming problems to identify the optimal solution of $x_1 = 0$, $x_2 = 0$, $x_3 = 8$, $x_4 = 7$. The optimal objective function value is -2380. Figure 5 shows the branch and bound tree, and Table 1 shows more detailed information about the various nodes on the tree. We chose $UB_X^0 = (0, 0, 0, 0)$ since x^0 was infeasible. A better solution was found at node 1, a still better one was found at node 5, and the optimal solution was found at node 8.

TABLE 1. *Solution data for the illustrative example*

Node number (k)	Branched from node	σ^k	$LB(N^k)$	x^k	UB_ϕ^k	UB_X^k
0	—	(0, 0, 0, 0)	−2,577.38	(0, 0, 7.14, 7.86)	0	(0, 0, 0, 0)
1	0	(0, 0, 0, 1)	−2,030.00	(0, 0, 15, 0)	−2,030	(0, 0, 15, 0)
2	0	(0, 0, 0, 2)	−1,774.00	(0, 0, 12, 3)	−2,030	(0, 0, 15, 0)
3	0	(0, 0, 0, 3)	−1,955.55	(0, 0, 10.42, 4.58)	−2,030	(0, 0, 15, 0)
4	0	(0, 0, 0, 4)	−2,480.95	(0, 0, 7.14, 7.86)	−2,030	(0, 0, 15, 0)
5	4	(0, 0, 1, 4)	−2,050.00	(0, 0, 0, 10)	−2,050	(0, 0, 0, 10)
6	4	(0, 0, 2, 4)	−2,050.00	(0, 0, 4, 8.8)	−2,050	(0, 0, 0, 10)
7	4	(0, 0, 3, 4)	−2,300.00	(0, 1.25, 6.25, 7.5)	−2,050	(0, 0, 0, 10)
8	4	(0, 0, 4, 4)	−2,380.00	(0, 0, 8, 7)	−2,380	(0, 0, 8, 7)
9	4	(0, 0, 5, 4)	$+\infty$	(0, 0, 8, 7)	−2,380	(0, 0, 8, 7)

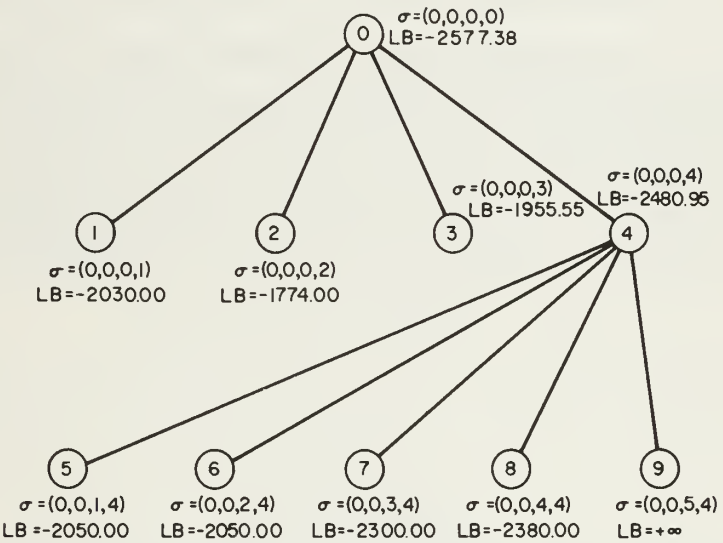


FIGURE 5. Branch and bound tree for the illustrative example.

REFERENCES

[1] Agin, N., "Optimum Seeking with Branch-and-Bound," Management Science *13*, B176–B185 (1966).

[2] Balas, E., "A Note on the Branch and Bound Principle," Operations Research *16*, 442–445 (1968).

[3] Beale, E. M. L., "Two Transportation Problems," in Preceedings of the Third International Conference on Operational Research, G. Krewas and G. Morlat (Eds.) pp. 780–788. (Dunod, Paris, 1964).

[4] Beale, E. M. L., *Mathematical Programming in Practice* (Pittman, London, 1968).

[5] Dorfman, R., P. Samuelson, and R. Solow, *Linear Programming and Economic Analysis* (McGraw-Hill, New York, 1958).

[6] Falk, J. E., "Lagrange Multipliers and Nonconvex Programs," SIAM Journal on Control 7, 534–545 (1969).

[7] Fiacco, A. V. and G. P. McCormick, *Nonlinear Programming: Sequential Uncontrained Minimization Techniques* (Wiley, New York, 1968).

[8] Jones, A. P. and R. M. Soland, "A Branch-and-Bound Algorithm for Multi-Level Fixed-Charge Problems," *Management Science 16*, 67–76 (1969).

[9] Mitra, G., "Dichotomizing Procedure for the Fixed Charge Transportation Problem," ICSI Report 105, Institute of Computer Science, University of London (Mar. 1968).

[10] Mitra, G., "Dichotomizing Procedure for the Fixed Charge Problem," ICSI Report 122, Institute of Computer Science, University of London (June 1968).

[11] Mitten, L. G., "Branch and Bound Methods: General Formulation and Properties," Operations Research *18*, 24–34 (1970).

[12] Soland, R. M., "An Algorithm for Separable Nonconvex Programming Problems II: Nonconvex Constraints," Management Science, *17*, 759–773 (1971).

HYPERBOLIC INTEGER PROGRAMMING*

M. Grunspan

U.S. Atomic Energy Commission

and

M. E. Thomas

University of Florida

ABSTRACT

The hyperbolic integer program is treated as a special case of a hyperbolic program with a finite number of feasible points. The continuous hyperbolic program also belongs to this class since its solution can be obtained by considering only the extreme points of the feasible set. A general algorithm for solving the hyperbolic integer program which reduces to solving a sequence of linear integer problems is proposed. When the integer restriction is removed, this algorithm is similar to the Isbell-Marlow procedure. The geometrical aspects of the hyperbolic problem are also discussed and several cutting plane algorithms are given.

1. THE HYPERBOLIC INTEGER PROGRAMMING PROBLEM

In addition to arising in a natural way [5], the hyperbolic programming model may also be used as an alternative to the classical linear programming model. A frequent example which occurs in economics is finding the most favorable ratio of revenues and allocations subject to restrictions on the availability of goods. When dealing with indivisible commodities, semblance to reality can be maintained by imposing an integer restriction on the components of the solution vector. This leads to the following problem:

$$(P') \qquad \max \{f(\mathbf{x}') = (\mathbf{c}'\mathbf{x}' + \alpha)/(\mathbf{d}' \quad + \beta)\}$$

subject to

$$A'\mathbf{x}' \leqslant \mathbf{b}$$

$$x' \geqslant 0 \text{ and integer.}$$

In (P'), A' is an $(m \times n)$ integer matrix, $\mathbf{x}'$ is an n dimensional integer column vector, $\mathbf{c}'$ and $\mathbf{d}'$ are n dimensional integer row vectors, $\mathbf{b}$ is an m dimensional integer column vector and α and β are fixed integers.

An alternate formulation for (P') is obtained by adding slack variables to the constraint equations. This leads to the following problem:

$$(P) \qquad \max \{f(\mathbf{x}) = (\mathbf{c}\mathbf{x} + \alpha)/(\mathbf{d}\mathbf{x} + \beta)\}$$

*Research for this report was sponsored, in part, by project THEMIS, ARO–D Contract No. DA HC04–68C–0002.

subject to

$$\mathbf{Ax} = \mathbf{b}$$

$$\mathbf{x} \geq 0 \text{ and integer.}$$

In (P), A is an $m \times (m+n)$ integer matrix, $\mathbf{x}$ is an $(m+n)$ dimensional integer column vector, $\mathbf{c}$ and $\mathbf{d}$ are $(m+n)$ dimensional integer row vectors, $\mathbf{b}$ is an m dimensional integer column vector, and α and β are fixed integers.

Problems (P′) and (P) are alternate statements of the general hyperbolic integer programming problem. Solving (P) is equivalent to solving (P′). The objective of this paper is to develop a class of algorithms for finding the optimal solution for (P) or (P′).

Throughout the remainder of this paper it will be assumed that the set of feasible solutions for (P) is bounded and that the value of the denominator in the objective function is strictly greater than zero for all feasible solutions. These assumptions are quite realistic in an economic context. Respectively they imply that the levels of activities being programmed cannot be unbounded and that the objective function value cannot become infinite.

2. SOLVING HYPERBOLIC PROGRAMS WITH A FINITE NUMBER OF FEASIBLE POINTS

The general algorithm discussed in this paper can be viewed as a generalization of an algorithm originally proposed by Isbell and Marlow [7] for solving the continuous hyperbolic program. It will be useful to first review the central idea behind this algorithm. Let $\overline{\mathbf{X}}' = (\mathbf{x}' | A'\mathbf{x}' \leq \mathbf{b}, \mathbf{x}' \geq 0)$ and let $\mathbf{x}'_a$ be any point such that $(\mathbf{d}'\mathbf{x}'_a + \beta) \neq 0$. In addition let (P'_C) denote the continuous hyperbolic programming problem. First suppose that the point $\mathbf{x}'_a$ is feasible for (P'_C) and consider the following linear programming problem:

$$\max\,(z(\mathbf{x}') = (\mathbf{d}'\mathbf{x}'_a + \beta)\,(\mathbf{c}'\mathbf{x}' + \alpha) - (\mathbf{c}'\mathbf{x}'_a + \alpha)\,(\mathbf{d}'\mathbf{x}' + \beta))$$

subject to

$$\mathbf{x}' \epsilon \overline{\mathbf{X}}'.$$

Since the point $\mathbf{x}'_a$ is feasible for this problem, it follows that $\max z(\mathbf{x}') \geq 0$. Thus if $\mathbf{x}'_1$ is an optimal solution for the linear program, the condition $z(\mathbf{x}'_1) \geq 0$ together with the assumption that

$$(\mathbf{d}'\mathbf{x}' + \beta) > 0 \,\forall\, \mathbf{x}' \epsilon \overline{\mathbf{X}}'$$

implies that $f(\mathbf{x}'_1) \geq f(\mathbf{x}'_a)$. Hence, given a feasible point $\mathbf{x}'_a$ for (P'_C), it is always possible to formulate a linear program whose solution yields an objective function value for (P'_C) which is at least as good as $f(\mathbf{x}'_a)$.

Interestingly enough, even if $\mathbf{x}'_a$ is not feasible for (P'_C), the solution to the linear program is still feasible for (P'_C). It is possible however that $f(\mathbf{x}'_1) \leq f(\mathbf{x}'_a)$. Since $\mathbf{x}'_a$ is not feasible, this is of no consequence. Of importance is that if a feasible point is not immediately available, one can always be obtained by solving a linear program. Once a feasible point is at hand, it is possible to formulate a second linear program which yields an objective function value for (P'_C) which is at least as good as the current value. Independent of the feasibility of $\mathbf{x}'_a$, the existence of a finite solution for the linear program is insured if $\overline{\mathbf{X}}'$ is assumed bounded.

The analysis given above suggests that the optimal solution for (P'_C) can be obtained by solving a sequence of linear programming problems. It is the basis for the Isbell-Marlow algorithm.

A Generalization

Consider the problem

$$(P'_F) \qquad \max\,(f(\mathbf{x}') = (\mathbf{c}'\mathbf{x}' + \alpha)/\,(\mathbf{d}'\mathbf{x}' + \beta))$$

subject to:

$$\mathbf{x}' \epsilon F,$$

where the feasible set F consists of a finite number of points. Following a line of reasoning given earlier, if $\mathbf{x}'_a$ is any feasible point for (P'_F), an improved solution for (P'_F) can be obtained by considering the following problem:

$$(LP'_F) \qquad \max\,(z(\mathbf{x}') = (\mathbf{d}'\mathbf{x}'_a + \beta)\,(\mathbf{c}'\mathbf{x}' + \alpha) - (\mathbf{c}'\mathbf{x}'_a + \alpha)\,(\mathbf{d}'\mathbf{x}' + \beta))$$

subject to

$$\mathbf{x}' \epsilon F.$$

Clearly, the ease with which a solution to (LP'_F) can be obtained depends on the structure of the set F.

The next theorem gives a necessary and sufficient condition for the point $\mathbf{x}'_a$ to be optimal for (P'_F).

THEOREM 1: The point $\mathbf{x}'_a \epsilon F$ maximizes the hyperbolic objective function in (P'_F) iff $\mathbf{x}'_a$ maximizes the objective function in (LP'_F).

PROOF: First suppose that

$$(1) \qquad \max f(\mathbf{x}') = f(\mathbf{x}'_a)$$

$$\mathbf{x}' \epsilon F.$$

Assuming that $(\mathbf{d}'\mathbf{x}' + \beta) > 0\ \forall \mathbf{x}' \epsilon F$, (1) implies that

$$(2) \qquad (\mathbf{d}'\mathbf{x}'_a + \beta)\,(\mathbf{c}'\mathbf{x}' + \alpha) - (\mathbf{c}'\mathbf{x}'_a + \alpha)\,(\mathbf{d}'\mathbf{x}' + \beta) \leq 0\ \forall \mathbf{x}' \epsilon F.$$

The left hand side in (2) is clearly the objective function in (LP'_F). By inspection, it achieves its upper bound at the point $\mathbf{x}'_a$. The point $\mathbf{x}'_a$ is therefore optimal for (LP'_F).

To prove the second part of the Theorem, suppose that

$$\max\,(z(\mathbf{x}')) = 0$$

subject to

$$\mathbf{x}' \epsilon F.$$

This implies that

$$z(\mathbf{x}') \leq 0 \ \forall \mathbf{x}' \epsilon F.$$

Since $(\mathbf{d}'\mathbf{x}' + \beta) > 0 \ \forall \mathbf{x}' \epsilon F$, the above inequality implies that

$$f(\mathbf{x}') \leq f(\mathbf{x}'_a) \ \forall \mathbf{x}' \epsilon F.$$

The point $\mathbf{x}'_a$ is therefore an optimal solution for (P'_F).

COROLLARY 1.1: The point $\mathbf{x}'_b$ which maximizes (LP'_F) is optimal for (P'_F) iff $z(\mathbf{x}'_b) = 0$.

PROOF: This result is a direct consequence of the proof given for Theorem 1.

Theorem 1 and Corollary 1.1 together with the observation that $\max_{\mathbf{x}' \epsilon F} z(\mathbf{x}') \geq 0$ suggests that the solution to (P'_F) can be obtained by solving a finite sequence of linear optimization problems.

A General Algorithm for the Hyperbolic Integer Program

Problem (P') with $\bar{\mathbf{X}}'$ bounded implies that the number of feasible integer points is finite. It is therefore possible to view (P') as a special case of (P'_F). Theorem 1 and Corollary 1.1 provide the basis for an algorithm to solve (P'). Such an algorithm is given below.

Algorithm HIP(1)

1. Find any feasible integer point belonging to $\bar{\mathbf{X}}'$. Call the point $\mathbf{x}'_0$ and set $j = 1$.
2. Define the following linear integer program:

$$\max \ (z_j = (\mathbf{d}'\mathbf{x}'_{j-1} + \beta)(\mathbf{c}'\mathbf{x}' + \alpha) - (\mathbf{c}'\mathbf{x}'_{j-1} + \alpha)(\mathbf{d}'\mathbf{x}' + \beta))$$

subject to

$$\mathbf{x}' \epsilon \bar{\mathbf{X}}'$$

$$\mathbf{x}' \text{ integer.}$$

3. Solve the problem defined in step 2 by any convenient method. Call the solution $\mathbf{x}'_j$.
4. If max $z_j = 0$, stop; otherwise let $j = j + 1$ and return to step 2.

The convergence of HIP(1) depends solely on having a finite number of feasible integer points to consider. For convenience, it is assumed that there are K feasible integer points. By construction, the algorithm generates a sequence of points $(\mathbf{x}'_i)$ such that $f(\mathbf{x}'_i) \leq f(\mathbf{x}'_{i+1}) \ \forall \ i$. If there exists some finite index such that $f(\mathbf{x}'_j) = f(\mathbf{x}'_{j+1})$, it follows that $z_{j+1}(\mathbf{x}'_{j+1}) = 0$. Appealing to Corollary 1.1, $\mathbf{x}'_j$ and $\mathbf{x}'_{j+1}$ are optimal hyperbolic solutions. In the worst case HIP(1) will generate all K feasible integer points such that $f(\mathbf{x}'_1) \leq f(\mathbf{x}'_2) \leq \ . \ . \ . \ \leq f(\mathbf{x}'_K)$. Because the functional values generated by HIP(1) are such that $f(\mathbf{x}'_i) \leq f(\mathbf{x}'_{i+1}) \ \forall \ i$, it must be that $f(\mathbf{x}'_K) = f(\mathbf{x}'_{K+1})$. This implies that $z_{K+1}(\mathbf{x}'_{K+1}) = 0$. Appealing to Corollary 1.1, $\mathbf{x}'_K$ and $\mathbf{x}'_{K+1}$ must be optimal hyperbolic solutions. The boundedness assumption on $\bar{\mathbf{X}}'$ insures that only a finite number of feasible integer points exist for (P'). Thus if HIP(1) is used to solve (P'), the optimal solution will be obtained in a finite number of steps.

Special cases of (P') have been considered by other authors. For example when the components of the solution vector $\mathbf{x}'$ are restricted to be (0–1) variables and the constraint equations are pseudo-

boolean expressions, HIP(1) reduces to an algorithm similar to the one proposed by Robillard and Florian [3].

HIP(1) can also be used to solve (P'_C), the continuous version of (P'). Because the feasible region for (P'_C) has a finite number of extreme points, and because the optimal solution to (P'_C) must occur on at least one of these points, finding the optimal solution for (P'_C) is equivalent to solving a hyperbolic program with a finite number of feasible points. That is, it suffices to maximize the objective function over the set of extreme points belonging to $\bar{\mathbf{X}}'$. (P'_C) can therefore be viewed as a special case of (P'_F) and HIP(1) will converge in a finite number of steps if it used to solve (P'_C). The resulting algorithm is in fact similar to the algorithm proposed by Isbell and Marlow [7].

With regard to integer programming, HIP(1) can be used to solve problems of the following form:

$$(P'_G) \qquad \max (f(\mathbf{x}') = (\mathbf{c}'\mathbf{x}' + \alpha)/(\mathbf{d}'\mathbf{x}' + \beta))$$

subject to

$$\mathbf{x}' \in \bar{\mathbf{X}}'$$

$$\mathbf{x}' \text{ integer,}$$

where $\bar{\mathbf{X}}'$ is an arbitrary bounded convex set with integer points in its interior. How easily (P'_G) can be solved when HIP(1) is used depends on the difficulty of the subproblems which the algorithm generates. In the case when the feasible region is defined by parabolic constraints, the method of Witzgall [10] can be used to solve the subproblems. If $\bar{\mathbf{X}}'$ is an arbitrary convex set with integer points in its interior, the cutting plane method of Kelley [8] can be used to solve the integer subproblems. HIP(1) can clearly be used to solve the general hyperbolic integer program. The procedure will always terminate in a finite number of steps as long as there are a finite number of feasible integer points.

3. GEOMETRICAL ASPECTS OF HYPERBOLIC PROGRAMMING

This section is devoted to the geometry of the hyperbolic program. The ideas developed here will also be useful for describing geometrically the algorithms proposed in this paper.

Corresponding to any sequence of hyperbolic functional values $(f(\mathbf{x}'_1), \ldots, f(\mathbf{x}'_n))$ such that $f(\mathbf{x}'_i) < f(\mathbf{x}'_{i+1})$ $(i = 1, \ldots, n-1)$ there is a unique sequence of hyperplanes $(T^0_1, \ldots, T^0_n)$, where

$$(3) \qquad T^0_i = (\mathbf{x}' \mid (\mathbf{d}'\mathbf{x}'_i + \beta)(\mathbf{c}'\mathbf{x}' + \alpha) - (\mathbf{c}'\mathbf{x}'_i + \beta)(\mathbf{d}'\mathbf{x}' + \beta) = 0)$$

The hyperplanes T^0_i can be seen to intersect about the solution set for the following set of equations:

$$(4) \qquad \mathbf{c}'\mathbf{x}' + \alpha = 0$$

$$\mathbf{d}'\mathbf{x}' + \beta = 0.$$

The solution set for (4) is also the set

$$\text{kern } (T) = (\mathbf{x}' \mid (\mathbf{c}'\mathbf{x}' + \alpha = 0, \mathbf{d}'\mathbf{x}' + \beta = 0, \mathbf{x} \in E^n),$$

which is the kernel of the map defined by

$$T(\mathbf{x}') = (\mathbf{c}'\mathbf{x}' + \alpha, \mathbf{d}'\mathbf{x}' + \beta) \qquad \mathbf{x}' \epsilon E^n.$$

The first component of the image of $\mathbf{x}'$ under the map T corresponds to the numerator and the second component corresponds to the denominator of the hyperbolic objective function.

The set kern (T) is used to develop a foundation for describing geometrically the behaviour of the hyperbolic objective function in E^n.

A trivial hyperbolic program occurs when the functionals $(\mathbf{c}'\mathbf{x}' + \alpha)$ and $(\mathbf{d}'\mathbf{x}' + \beta)$ are linearly dependent. This occurrence leads to a constant objective function value $\forall \mathbf{x}' \epsilon E^n$. In this discussion it will be assumed that $(\mathbf{c}'\mathbf{x}' + \alpha)$ and $(\mathbf{d}'\mathbf{x}' + \beta)$ are linearly independent. This assumption leads to the following result:

LEMMA 1: The set kern $(T) \neq \Phi$ if the functional $(\mathbf{c}'\mathbf{x}' + \alpha)$ is linearly independent of the functional $(\mathbf{d}'\mathbf{x}' + \beta)$.

PROOF: The linear independence of $(\mathbf{c}'\mathbf{x}' + \alpha)$ and $(\mathbf{d}'\mathbf{x}' + \beta)$ implies that the map T is onto. Hence there exists some point $\mathbf{x}'_*$ which T maps into the point (0, 0) and kern $(T) \neq \Phi$.

Based on the assumptions made thus far, kern $(T) \neq \Phi$ and $\overline{\mathbf{X}}' \neq \Phi$. With the additional assumption that $(\mathbf{d}'\mathbf{x}' + \beta) > 0$, it also follows that kern $(T) \cap \overline{\mathbf{X}}' = \Phi$. The next Lemma provides the tool for showing that a change in the hyperbolic objective function value can be achieved by a rotation of a hyperplane T_i^0 about kern (T).

LEMMA 2: When $(\mathbf{d}'\mathbf{x}' + \beta) > 0 \ \forall \mathbf{x}' \in \overline{\mathbf{X}}'$, the set kern (T) together with any $\mathbf{x}'_a \in \overline{\mathbf{X}}'$ define a hyperplane T_a^0 in E^n.

PROOF: If T is a linear map such that $T: E^n \rightarrow E^m$, kern (T) + dim range $(T) = n$. In particular when $m = 2$, kern (T) is an $(n-2)$ dimensional subspace of E^n. With the assumptions made thus far, kern $(T) \cap \overline{\mathbf{X}}' = \Phi$. Hence any point $\mathbf{x}'_a \in \overline{\mathbf{X}}'$ is linearly independent of any point belonging to kern (T) and any basis for kern (T) together with $\mathbf{x}'_a$ span an $(n-1)$ dimensional subspace of E^n. By definition, this is a hyperplane in E^n. Recalling the definition of T_a^0 given in (3), it is clear that in addition to the point $\mathbf{x}'_a$ every point in the set kern (T) lies on the hyperplane T_a^0. Hence kern (T) together with $\mathbf{x}'_a \in \overline{\mathbf{X}}'$ define T_a^0. This proves the Lemma.

To see that a change in the hyperbolic objective function value can be achieved by a rotation of T_i^0 about kern (T) recall that given a sequence $(f(\mathbf{x}'_1), \ldots, f(\mathbf{x}'_n))$ with $\mathbf{x}'_i \epsilon \overline{\mathbf{X}}' \ \forall i$ and $f(\mathbf{x}'_i) < f(\mathbf{x}'_{i+1})$ $(i = 1, \ldots, n-1)$ there exist a unique corresponding sequence $(T_1^0, \ldots, T_n^0)$. Appealing to Lemma 2, note that $\mathbf{x}'_i$ together with kern (T) define T_i^0 and $\mathbf{x}'_{i+1}$ together with kern (T) define T_{i+1}^0. Furthermore, by definition of the hyperplane T_i^0,

$$T_i^0 \cap T_{i+1}^0 = \text{kern}(T).$$

Since the functional values $f(\mathbf{x}'_i)$ and $f(\mathbf{x}'_{i+1})$ are achieved with points that lie on interesecting hyperplanes, and since it was originally assumed that $\mathbf{x}'_i \in \overline{\mathbf{X}}' \ \forall i$ and $f(\mathbf{x}'_i) < f(\mathbf{x}'_{i+1})$, the desired conclusion follows.

Since an increase in the hyperbolic objective value corresponds to a rotation of some hyperplane T_i^0 about kern (T), it is clear that the largest feasible value will be attained on at least one extreme point of $\overline{\mathbf{X}}'$. This result was proved by Martos [9] and Dorn [2] without relying on the geometrical aspects of the problem. This same result is also a consequence of Isbell-Marlow [7] algorithm.

Before proceeding, it will be useful to examine the Isbell-Marlow algorithm within the framework developed here. First a point $\mathbf{x}'_a$ such that $(\mathbf{d}'\mathbf{x}'_a+\beta)\neq 0$ is chosen and a feasible point $\mathbf{x}'_1$ which maximizes the expression

$$(\mathbf{d}'\mathbf{x}'_a+\beta)(\mathbf{c}'\mathbf{x}'+\alpha)-(\mathbf{c}'\mathbf{x}'_a+\alpha)(\mathbf{d}'\mathbf{x}'+\beta)$$

is obtained by solving a linear programming problem. If the hyperplane T^0_1 is a supporting hyperplane to the feasible region, the point $\mathbf{x}'_1$ is optimal for (P'_C). If $\mathbf{x}'_1$ is not optimal, a point $\mathbf{x}'_2$ which maximizes the expression

$$(\mathbf{d}'\mathbf{x}'_1+\beta)(\mathbf{c}'\mathbf{x}'+\alpha)-(\mathbf{c}'\mathbf{x}'_1+\alpha)(\mathbf{d}'\mathbf{x}'+\beta)$$

is obtained. If T^0_2 is a supporting hyperplane to the feasible region, $\mathbf{x}'_2$ is optimal. If it is not, the procedure is repeated until the optimal solution is recognized.

The most immediate way of obtaining an optimal solution for (P'_C) is to rotate some hyperplane T^0_a about kern (t). Although the Isbell-Marlow algorithm does not implement a rotation of T^0_a, it accomplishes the same end. It sweeps across the feasible region in a direction towards the optimal extreme point. Geometrically, the algorithm translates hyperplanes of changing slopes until the optimal solution for (P'_C) is recognized. The general algorithm HIP(1) can be viewed in a similar fashion. Optimality is recognized however when $T^0_i=T^0_{i+1}$.

4. CUTTING PLANE ALGORITHMS

Cutting plane methods for solving the linear integer programming problem were suggested by Danzig, Fulkerson, and Johnson [1] and formalized by Gomory [4]. The principle behind these methods is to reduce the original feasible region without eliminating any feasible integer points and then to reoptimize the objective function on the reduced feasible set. If the reoptimization phase yields an all integer solution, it is also optimal. Gomory accomplishes this by generating new constraints from existing constraints and then reoptimizes the objective function with the aid of the dual simplex algorithm. A desirable characteristic belonging to Gomory cuts is that they can be generated systematically and without much effort. Another important property of a Gomory cut is that it reduces the current feasible region by cutting off the current noninteger optimal extreme point so that no feasible integer points are eliminated. The fact that the optimal solution for (P_C) must lie on at least one extreme point of $\overline{\mathbf{X}}$ (see Martos [9] and Dorn [2]), suggests that Gomory cuts can also be used to solve (P). This observation is verified in the next section.

A Fractional Cutting Plane Algorithm

The algorithm described here is a synthesis of the Martos [9] algorithm, the Isbell-Marlow [7] algorithm, Gomory cutting planes and Lemke's dual simplex method for linear programs. Alternately, the algorithm can also be viewed as a special case of the general algorithm HIP(1) described in section 2.

The first step is to choose a feasible point, not necessarily an integer point, and construct a linear integer program whose solution is feasible for (P). In this algorithm, the feasible point is the optimal solution for (P_C) and it is obtained by using the column tableau for Martos's algorithm (see Table I). If the solution displayed in the optimal tableau is all integer, it is also optimal for (P). If it is not, a Gomory fractional cutting plane can be generated from the optimal representation of the constraint

TABLE I

		t_1	t_2	. . .	t_s	. . . .	t_n
	1	$-x_1$	$-x_2$	. . .	$-x_s$	. . . .	$-x_n$
N	n_{00}	n_{01}	n_{02}	. . .	n_{0s}	. . . .	n_{0n}
D	d_{00}	d_{01}	d_{02}	. . .	d_{0s}	. . . .	d_{0n}
z	a_{00}	a_{01}	a_{02}	. . .	a_{0s}	. . . .	a_{0n}
x_1	a_{10}	a_{11}	a_{12}	. . .	a_{1s}	. . . .	a_{1n}
.	.	.	.		.		.
.	.	.	.		.		.
.	.	.	.		.		.
.	.	.	.		.		.
x_n	a_{n0}	a_{n1}	a_{n2}	. . .	a_{ns}	. . . .	a_{nn}
x_{n+1}	$a_{n+1,0}$	$a_{n+1,1}$	$a_{n+1,2}$	. . .	$a_{n+1,s}$	. . . .	$a_{n+1,n}$
.	.	.	.		.		.
.	.	.	.		.		.
.	.	.	.		.		.
.	.	.	.		.		.
x_{n+m}	$a_{n+m,0}$	$a_{n+m,1}$	$a_{n+m,2}$	. . .	$a_{n+m,s}$	. . . .	$a_{n+m,n}$
s_i	$-f_{i0}$	$-f_{i1}$	$-f_{i2}$	. . .	$-f_{is}$	. . . .	$-f_{in}$

set. Appending a legitimate cutting plane to the optimal hyperbolic tableau does not exclude any feasible integer points and thus a reoptimization of the enlarged tableau can lead to the optimal integer solution. Unfortunately, the Gomory fractional cutting plane destroys the primal tableau and there is no analog of the dual simplex method for the hyperbolic program which can be used to reoptimize the tableau. Primal feasibility can however be regained by solving

(P1) $$\max(z=(\mathbf{dx}_*+\beta)(\mathbf{cx}+\alpha)-(\mathbf{cx}_*+\alpha)(\mathbf{dx}+\beta))$$

subject to

$$A\mathbf{x}=\mathbf{b}$$

$$\mathbf{x}\geqslant 0 \text{ and integer},$$

where $\mathbf{x}_*$ is the extreme point of the feasible set which maximizes the continuous hyperbolic program. (P1) is solved using Gomory's fractional algorithm. As a result $\mathbf{x}_*$ is eliminated from the feasible region and furthermore no feasible integer points are lost. Solving (P1) requires that the optimal tableau without the integer restriction be obtained first. It is easily verified that this tableau is available without any additional work if the algorithm of Martos with the column tableau is used to solve (P_C). Two tests are available to determine if the solution to (PI) is also optimal for (P). First, the relative hyperbolic costs which are denoted by t_j can be computed directly from Table I. If $t_j \leqslant 0$ for all the nonbasic vari-

ables, the solution to (P1) is also optimal for (P). As a second test, if the optimal objective function value for (P1) is exactly zero, the solution to (P1) is also optimal for (P) (see step 4 in HIP(1)).

Let $\mathbf{x}_1$ denote the solution to (P1) and define S_1 to be the region generated by the Gomory cuts used to solve (P1). If after $\mathbf{x}_1$ is obtained, there exists $t_j \geq 0$ for some nonbasic variable, it is still possible for the current solution to be optimal for (P). To see if this is indeed the case, the following problem is formulated

(P2) $$\max(z = (\mathbf{dx}_1 + \beta)(\mathbf{cx} + \alpha) - (\mathbf{cx}_1 + \alpha)(\mathbf{dx} + \beta))$$

subject to

$$\mathbf{x} \in \overline{\mathbf{X}} \cap S_1$$

$$\mathbf{x} \text{ integer.}$$

The optimal solution to (P2) is again obtained using Gomory's algorithm. The noninteger solution which is obtained first is called $\mathbf{x}_1^*$. Note that the hyperplane $z(\mathbf{x}_1^*) = (\mathbf{dx}_1 + \beta)(\mathbf{cx}_1^* + \alpha) - (\mathbf{cx}_1 + \alpha)(\mathbf{dx}_1^* + \beta)$ is parallel to T_1^0. Thus if a cutting plane algorithm is used to solve (P2), a new integer extreme point on the integer hull contained in the original feasible region is obtained or an integer point lying on T_1^0 is obtained. If the latter occurs, $z(\mathbf{x}_2) = 0$ and $\mathbf{x}_2$ which is the optimal solution to (P2) is also optimal for (P). If the solution to (P2) does not lie on T_1^0 and there exists some $t_j > 0$, a new linear integer program is defined and the procedure is repeated. Before giving a statement of a fractional cutting plane algorithm, it will be useful to define the following notation:

$\overline{\mathbf{X}}_1 = \overline{\mathbf{X}} = (x \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0)$
$\overline{\mathbf{X}}_j$ = feasible region for jth linear integer program
S_j = region defined by Gomory cuts used to solve jth linear integer program

Algorithm HIP(2)

1. Disregard the integer restriction and solve (P) by use of the column tableau version of the Martos algorithm. Call the solution $\mathbf{x}_*$. If the solution is all integer stop; otherwise let $j = 1$ and $\mathbf{x}_0 = \mathbf{x}_*$.

2. Define the following linear integer program:

$$\max\ (z_j = (\mathbf{dx}_{j-1} + \beta)(\mathbf{cx} + \alpha) - (\mathbf{cx}_{j-1} + \alpha)(\mathbf{dx} + \beta))$$

subject to

$$\mathbf{x} \in \overline{\mathbf{X}}_j$$

$$\mathbf{x} \text{ integer,}$$

where

$$\overline{\mathbf{X}}_1 = \overline{\mathbf{X}}$$

$$\overline{\mathbf{X}}_j = \overline{\mathbf{X}}_{j-1} \cap S_{j-1} \qquad \text{for } j = 2, 3, \ldots$$

3. Solve the problem is step 2 using Gomory's fractional algorithm. Call the solution $\mathbf{x}_j$.
4. If max $z_j=0$ or if $t_j \leq 0\ \forall j$, stop; otherwise let $j=j+1$ and return to step 2.

Although HIP(2) is clearly a special case of HIP(1), it is also in the same spirit as Gomory's fractional algorithm for the linear integer program. That is, the optimal extreme point for (P_C) is obtained first. If the solution is integer, the procedure terminates. If it is not, a cutting plane which cuts off the current extreme point and which does not delete any feasible integer points is appended to the original feasible region. The objective function is then reoptimized on the reduced feasible region. New cutting planes are generated until the reoptimization phase yields an extreme point of the integer hull contained in $\bar{X}$ What makes (P) more difficult to solve than the linear integer program is that more than one extreme point of the integer hull may have to be generated before the optimality condition is satisfied. Since $\bar{X}$ is assumed bounded, the integer hull has a finite number of extreme points. At most HIP(2) will generate all of these and since $f(\mathbf{x}_{j+1}) \geq f(\mathbf{x}_j)\ \forall j$, HIP(2) will converge in a finite number of steps.

Although conceptually the same, HIP(1) and HIP(2) do have some differences. In HIP(2), the feasible region is reduced at each iteration whereas in HIP(1) the feasible region remains the same at each iteration. Because of the nature of cutting plane algorithms, HIP(2) only searches points which lie within the boundary of the original feasible region and which are outside or on the boundary of the integer hull contained in the original feasible region. This is not necessarily the case with HIP(1) if a branch and bound algorithm is used to solve each of the linear integer programs. A special property of HIP(2) is that the optimal tableau for the jth linear integer problem with some slight modification becomes the initial tableau for the $(j+1)$th problem.This simply means that when solving the $(j+1)$th problem, it is not necessary to start from the beginning.

The cutting plane algorithm described here is implemented with the aid of the tableau format shown in Table I. The entries in the first row are the relative costs t_j which are used to check optimality. They are computed after each integer solution is obtained according to

$$t_j = n_{00}d_{0j} - n_{0j}d_{00} \qquad \forall j. \tag{5}$$

The entries in the second row identify the nonbasic variables at the current iteration. The N and D rows represent the numerator and denominator, respectively. The entry n_{00} is the current numerator value and d_{00} is the current denominator value. The z row denotes the current linear objective whose integer solution is being sought. Rows x_{n+1} through x_{n+m} are the current basic variables. The last row is reserved for the Gomory cut. To demonstrate the workings of HIP(2), consider the following example:

$$\max f(x_1, x_2) = (5x_1+3x_2+1)/(5x_1+2x_2+1)$$

subject to

$$3x_1+5x_2 \leq 16$$

$$5x_1+2x_2 \leq 11$$

$$x_1, x_2 \geq 0 \text{ and integer.}$$

Slack variables x_3 and x_4 are introduced and the constraints are rewritten as

$$x_3 = 16-3x_1-5x_2$$

$$x_4 = 11 - 5x_1 - 2x_2.$$

Table II is a statement of the original problem and Table III exhibits the optimal continuous solution. In Table IV, the proper z row and a Gomory cut generated from the x_2 row are appended to the tableau. Pivoting results in Table V. Upon calculating the t_j's, it is found that the current solution is also optimal for the original problem.

TABLE II

t_j		0	1
	1	$-x_1$	$-x_2$
N	1	−5	−3
D	1	−5	−2
x_1	0	−1	0
x_2	0	0	−1
x_3	16	3	5
x_4	11	5	2

TABLE III

t_j		−83/5	−1/5
	1	$-x_1$	$-x_3$
N	53/5	−16/5	3/5
D	37/5	−19/5	2/5
x_1	0	−1	0
x_2	16/5	3/5	1/5
x_3	0	0	−1
x_4	23/5	19/5	−2/5

TABLE IV

t_j		−83/5	−1/5
	1	$-x_1$	$-x_3$
N	53/5	−16/5	3/5
D	37/5	−19/5	2/5
z	0	83/5	1/5
x_1	0	−1	0
x_2	16/5	3/5	1/5
x_3	0	0	−1
x_4	23/5	19/5	−2/5
s_1	−1/5	−3/5	−1/5

TABLE V

t_j		−15	−1
	1	$-x_1$	$-s_1$
N	10	−5	3
D	7	−5	2
z	−1	16	1
x_1	0	−1	0
x_2	3	0	1
x_3	1	3	−5
x_4	5	5	−2
s_1	0	0	−1

An All Integer Cutting Plane Algorithm

The algorithm discussed in this section is based on the Gomory all integer cut. It is also a special case of HIP(1) and it is initiated by displaying the problem in column tableau format. The entries in the z-row for the first linear problem are $a_{00} = 0$ and $a_{0j} = -t_j$, where t_j is defined in (5). If $a_{0j} \geq 0 \; \forall j$, increasing any one of the original variables to positive levels will decrease the current objective function value. Therefore, if $a_{0j} \geq 0 \; \forall j$ the optimal objective function value is (α/β). If there exist some $a_{0j} < 0$

the initial tableau is primal feasible with respect to the relative costs displayed in the z-row. It is not, however, dual feasible. As in HIP(2), the first step is to find a feasible integer point for (P). In addition an all integer tableau is to be maintained after each pivot. Each of these objectives is achieved if the linear function displayed in the z-row is maximized using either a primal cutting plane algorithm or Gomory's all integer algorithm. If a primal cutting plane algorithm is used, the first pivot can be made without altering the first tableau. If the Gomory all integer algorithm is used, the current tableau must first be rendered dual feasible with respect to the relative costs of the linear objective function. This is accomplished by appending the constraint

$$x_{n+m+1} = M - \Sigma x_j, \tag{6}$$

to the current tableau and choosing it as the pivot row. The pivot column is chosen as the lexicographically smallest column with respect to the current linear program. The constant M in (6) is an arbitrarily large integer such that $x_{n+m+1} \geq 0$. The constraint given by (6) simply expresses the fact that the original constraint set is bounded. Pivoting at this point will result in an all integer dual feasible tableau. A statement of the all integer algorithm is given below.

Algorithm HIP(3)

1. Set up the column tableau for the hyperbolic program as in Table I.

The components of the vector $\mathbf{x}_0$ are $x_{i0} = 0$ $(i = 1, \ldots, n)$ and $x_{n+j,0} = b_j$ $(J = 1, \ldots, m)$. Set $j = 1$.

2. Generate the following linear integer program:

$$\max [z_j = (\mathbf{d}\mathbf{x}_{j-1} + \beta)(\mathbf{c}\mathbf{x} + \alpha) - (\mathbf{c}\mathbf{x}_{j-1} + \alpha)(\mathbf{d}\mathbf{x} + \beta)]$$

subject to

$$\mathbf{x} \epsilon \overline{\mathbf{X}}_j$$

$$\mathbf{x} \text{ integer},$$

where

$$\overline{\mathbf{X}}_1 = \overline{\mathbf{X}}$$

$$\overline{\mathbf{X}}_j = \overline{\mathbf{X}}_{j-1} \cap S_{j-1} \qquad j = 2, 3, \ldots$$

3. *Option A:* Solve the problem defined in step 2 using a primal cutting plane algorithm.

Option B: Append the constraint given by (6) and pivot to render tableau dual feasible with respect to the relative costs in the z-row. Use the Gomory all integer algorithm to maximize the linear objective function.

4. Compute $t_j \forall_j$. If $t_j \leq 0 \forall_j$ or if $z_j(\mathbf{x}_j) = 0$ stop; otherwise let $j = j + 1$ and return to step 2.

Since HIP(3) is a special case of HIP(1), the finite convergence proofs are the same. It should be pointed out that when implementing HIP(3) with option B, the constant M in (6) must be chosen with some care. It is important to realize that a constraint of the form (6) must be appended every time a linear integer program is generated. Since the set of nonbasic variables changes for every linear integer program, a constant M which is large enough for the first problem may not be large enough for the

second problem. This simply means that the constant may have to be changed at each iteration. Certainly, if it is chosen large enough at the outset, it need never be changed. The example given in section 7 is also used to demonstrate HIP(3).

TABLE VI

t_j		0	1
	1	$-x_1$	$-x_2$
N	1	−5	−3
D	1	−5	−2
z	0	0	−1
x_1	0	−1	0
x_2	0	0	−1
x_3	16	3	5
x_4	11	5	2
x_5	20	1	1

Table VI is a statement of the original problem with the z-row included. An additional row exhibiting the constraint

$$x_5 = 20 - x_1 - x_2,$$

TABLE VII

t_j			
	1	$-x_1$	$-x_5$
N	61	−5	3
D	41	−5	2
z	20	0	1
x_1	0	−1	0
x_2	20	0	1
x_3	−84	3	−5
x_4	−29	5	−2
x_5	0	1	−1
s_1	−17	0	−1

TABLE VIII

t_j		−15	−1
	1	$-x_1$	$-s_1$
N	10	−5	3
D	7	−5	2
z	3	0	1
x_1	0	−1	0
x_2	3	0	1
x_3	1	3	−5
x_4	5	5	−2
x_5	17	1	−1
s_1	0	0	−1

where 20 is an upper bound on the sum of the current nonbasic variables is also appended. Table VII is the dual feasible tableau with respect to the z-row which results from the first pivot operation. The x_3 row is the source row and s_1 is the first cut appended to the original set of constraints. Reoptimizing z after s_1 is appended yields the optimal solution which is displayed in Table VIII.

5. A VARIANT OF THE FRACTIONAL CUTTING PLANE ALGORITHM

This algorithm is initiated in precisely the same way as HIP(2). The counter index is set equal to one and the first three steps are repeated. If $t_j \leq 0 \forall_j$, the surrent solution is also optimal for the hyperbolic program. If there exist $t_j > 0$, this algorithm makes use of the fact that $f(\mathbf{x}_j)$ is a lower bound for the maximum of the hyperbolic objective function. In other words, the integer solution must satisfy the following constraint:

$$(\mathbf{dx}_j + \beta)(\mathbf{cx} + \alpha) - (\mathbf{cx}_j + \alpha)(\mathbf{dx} + \beta) \geq 0. \tag{7}$$

The strategy in this algorithm is to append (7) to the current tableau and use the Martos [9] algorithm to reoptimize the hyperbolic objective function. This step yields the optimal hyperbolic extreme point for the current feasible region. If the optimal solution has all integer components, the algorithm terminates. If it does not, let $\mathbf{x}_j^*$ denote the optimal hyperbolic noninteger extreme point and generate the following linear integer program:

$$\max\ (z_{j+1} = (\mathbf{dx}_j^* + \beta)(\mathbf{cx} + \alpha) - (\mathbf{cx}_j^* + \alpha)(\mathbf{dx} + \beta))$$

subject to

$$\mathbf{x} \in \bar{\mathbf{X}}_{j+1}$$

$$\mathbf{x} \text{ integer},$$

where T_j^0 is defined in (3), T_j is (7) with strict inequality, and $\bar{\mathbf{X}}_{j+1} = \bar{\mathbf{X}}_j \cap S_j \cap (T_j \, UT_j^0)$ $(j = 1, 2, \ldots)$. The solution to this problem is either optimal for the hyperbolic integer program or it is an improved lower bound. The algorithm is described below.

Algorithm HIP(4)

1. Solve (P) disregarding the integer restriction. If the solution is all integer, stop; otherwise set $\mathbf{x}_0^* = \mathbf{x}^*$ and $j = 1$.

2. Define the following linear integer program:

$$\max\ z_j = (\mathbf{dx}_{j-1} + \beta)(\mathbf{cx} + \alpha) - (\mathbf{cx}_{j-1} + \alpha)(\mathbf{dx} + \beta)$$

subject to

$$\mathbf{x} \in \bar{\mathbf{X}}_j, \ \mathbf{x} \text{ integer},$$

where

$$\bar{\mathbf{X}}_1 = \bar{\mathbf{X}}$$

$$\bar{\mathbf{X}}_j = \bar{\mathbf{X}}_{j-1} \cap S_{j-1} \cap (T_{j-1} \cup T_{j-1}^0) \qquad j = 2, 3, \ldots$$

3. Solve the problem defined in step 2 using Gomory's fractional algorithm. Call the solution $\mathbf{x}_j$.

4. If $t_j \leq 0 \; \forall j$ or if $f(\mathbf{x}_{j-1}) = f(\mathbf{x}_j)$ stop; otherwise go to step 5.

5. Use the Martos algorithm to reoptimize the current tableau. If the resulting solution is integer, stop; otherwise let $\mathbf{x}_j^*$ be the optimal extreme point and let $j = j+1$ and return to step 2.

The feasible region $\bar{\mathbf{X}}_{j+1}$ is defined by appending the constraint given in (7) to the optimal tableau for the $(j+1)$th problem and is discarded thereafter. The reason for this is that when

$$(\mathbf{d}\mathbf{x}_{j+1}+\beta)(\mathbf{c}\mathbf{x}+\alpha) - (\mathbf{c}\mathbf{x}_{j+1}+\alpha)(\mathbf{d}\mathbf{x}+\beta) \geq 0$$

is appended to $\mathbf{X}_{j+1} \cap S_{j+1}$, the result is that $\bar{\mathbf{X}}_{j+2} \subset \bar{\mathbf{X}}_{j+1}$.

The algorithm HIP(4) can be helpful when it is desirable to terminate at some near optimal integer point. Note that the quantities $f(\mathbf{x}_j)$ and $f(\mathbf{x}_j^*)$ are, respectively, the lower and upper bounds for the maximum of the hyperbolic function. Thus the quantity $\delta = f(\mathbf{x}_j^*) - f(\mathbf{x}_j) \geq 0$ is a measure of how close $\mathbf{x}_j$ is to the optimal integer solution. When $\mathbf{x}_j$ is optimal $\delta = 0$. A stopping rule is therefore to compute δ and stop when $\delta \leq \delta^*$, where δ^* is user specified.

HIP(4) converges in a finite number of steps because $\bar{\mathbf{X}}_{j+1} \subset \bar{\mathbf{X}}_j \; \forall j$ and the number of feasible integer points contained in $\bar{\mathbf{X}}_{j+1}$ is at least one less than the number of feasible integer points in $\bar{\mathbf{X}}_j$. Hence in the worst case there exists some finite iteration K such that $\bar{\mathbf{X}}_K$ contains exactly one integer point. Since by virtue of HIP(4) $f(\mathbf{x}_{j+1}) \geq f(\mathbf{x}_j)$, it follows that when there is only one integer point left in the feasible region $f(\mathbf{x}_{j+1}) = f(\mathbf{x}_j)$. Optimality is therefore achieved in a finite number of steps.

6. DISCUSSION

It is well known that cutting plane methods for solving integer programs have been successful only for specialized structures, and, in general, they are not computationally attractive. Thus algorithms HIP(2) and HIP(3) are primarily of theoretical interest. From a user's point of view, algorithm HIP(4) appears more promising since it provides an upper and a lower bound for the objective function at each iteration. This would, of course, permit cutting off the algorithm at some near optimal point if computation time becomes excessive.

In conclusion, the algorithms proposed here reduce to solving a sequence of linear integer programs. When the number of feasible integer points is finite, the hyperbolic algorithm is also finite. Clearly, the success of the algorithm depends on the state of the art of linear integer programming. Group theoretic methods can also be adapted to solve the hyperbolic integer program and have been studied by Grunspan in [6].

7. ACKNOWLEDGMENT

The authors wish to thank Dr. H. D. Ratliff of the University of Florida for his many helpful suggestions and comments.

REFERENCES

[1] Dantzig, G. B., D. R. Fulkerson, and S. M. Johnson, "Solution of a Large Scale Traveling-Salesman," Operations Research, *2*, 393–410 (1954).

[2] Dorn, W. S., "Linear Fractional Programming," IBM Research Rept. RC–830 (1962).

[3] Florian, M. and P. Robillard, "Hyperbolic Programming with Bivalent Variables," Department d'Informatique, Universite de Montreal, Publication #41 (Aug. 1970).

[4] Gomory, R. E., "An Algorithm for Integer Solutions to Linear Programs," in *Recent Advances in Mathematical Programming*, R. L. Graves and P. Wolfe (Eds.) pp. 269–302 (McGraw-Hill, New York, 1963).

[5] Grunspan, M., "Fractional Programming: A Survey," Technical Report No. 50, Project THEMIS, Systems Research Center, Industrial and Systems Engineering Department, University of Florida (Jan. 1971).

[6] Grunspan, M., "On the Integer Solution of the Hyperbolic Programming Problem," Technical Report No. 61, Project THEMIS, Systems Research Center, Industrial and Systems Engineering Department, University of Florida (Feb. 1972).

[7] Isbell, J. R. and W. H. Marlow, "Attrition Games," Nav. Res. Log. Quart. *3*, 71–94 (1965).

[8] Kelley, J. E., Jr., "The Cutting Plane Method for Solving Convex Programs," J. SIAM *8*, 703–712 (1960).

[9] Martos, B., "Hyperbolic Programming," Nav. Res. Log. Quart. *11*, 135–155 (1964).

[10] Witzgall, C., "An All-Integer Programming Algorithm with Parabolic Constraints," J. SIAM, 11, No. 4 (1963).

GEOMETRY AND RESOLUTION OF DUALITY GAPS

M. S. Bazaraa

Georgia Institute of Technology

ABSTRACT

In this study we interpret the exterior penalty function method as a generalized lagrangian method which fills duality gaps in nonconvex problems. Geometry and resolution of these gaps from a duality point of view are highlighted.

1. INTRODUCTION

Consider the following two problems. Problems P and D' are usually referred to as primal and dual problems. Here $f, g = (g_1, \ldots, g_m)$, and $h = (h_1, \ldots, h_k)$ are defined on R_n and S is a subset of R_n.

Problem P

Find $\inf \{f(x) : x \in S, g(x) \leq 0, h(x) = 0\}$.

Problem D'

Find $\sup \{\phi(\lambda, \mu) : \lambda \geq 0\}$, where $\phi(\lambda, \mu) = \inf \{f(x) + < \lambda, g(x) > + < \mu, h(x) > : x \in S\}$.

It is well known that $\inf \{f(x) : x \in S, g(x) \leq 0, h(x) = 0\} \geq \sup \{\phi(\lambda, \mu) : \lambda \geq 0\}$. Under suitable convexity assumptions and a suitable constraint qualification it is true that the above inequality holds as an equality, i.e., no duality gaps are encountered. Furthermore, x_0, λ_0, and μ_0 jointly solve problems P and D' above if, and only if, $\psi(x_0, \lambda, \mu) \leq \psi(x_0, \lambda_0, \mu_0) \leq \psi(x, \lambda_0, \mu_0)$ for each $x \in S$ and each $\lambda \geq 0$ and μ, where $\psi(x, \lambda, \mu) = f(x) + < \lambda, g(x) > + < \mu, h(x) >$. The reader may refer to Falk [6], Geoffrion [11], Rockafeller [16], and Bazaraa [1] for further details.

Now consider the problem: minimize $\{\psi(x, \lambda, \mu) : x \in S\}$ for fixed $\lambda \geq 0$ and μ. Denote the optimal solution by x'. From the above inequality, as Everett [5] noted, it is clear that x' will also solve the problem: minimize $\{f(x) : x \in S, g(x) \leq g(x'), h(x) = h(x')\}$. In particular if $g(x') = 0$ and $h(x') = 0$ then x' will solve the original problem P. Brooks and Geoffrion [4] strengthened Everett's results by showing that x' solves P provided that $g_i(x') = 0$ for each $i \in I$ and $g_i(x') \leq 0$ for each $i \in I$, where $I = \{i : \lambda_i = 0\}$. There are two main disadvantages for this approach, however. The first is that one is not able to find the $\lambda \geq 0$ and μ which work beforehand and a search procedure is required to do this (see [4]). The second and more serious disadvantage, is that such $\lambda \geq 0$ and μ may not even exist. This is the phenomenon usually referred to as a duality gap. See for example Everett [5] and Nunn [14] for discussion and partial solution to the gap problem Actually one is assured that such λ and μ exist only under the convexity assumptions of f, g, and S, the linearity of h, and under a suitable constraint qualification (see for example [11] and [1]). This means that the gap problem may very well occur for nonconvex problems and, in general, one may be unable to solve problem P by the above procedure.

In a significant contribution to the theory of lagrangian multipliers; Gould [12] showed that one may overcome the duality gap by using nonlinear multiplier functions. Gould showed the existence of multiplier functions which can "dip" into the gap regions. This ingenious idea, in theory enables one to get around the gap problem. However, in order to make use of this theory from an algorithmic point of view, one would like to identify suitable forms of nonlinear supports for given problems. At this stage, such an identification is not available. Addressing the same problem from the point of view of penalty functions (parametric supports), Bellmore, Greenberg, and Jarvis [3], discussed the problems associated with this approach, such as the problems of the existence of the nonlinear support, finding a supporting function (provided that it exists), and searching for the lagrangian multipliers. One may also refer to Greenberg [13] for detailed discussion on the gap problem and its resolution.

The reader should also refer to the significant work of Fiacco and McCormick [8–10] and Zangwill [17, 18] on penalty functions where problem P is converted into a sequence of unconstrained problems. Under suitable assumptions it is shown that the optimal solutions to this sequence converges to the optimal solution of problem P described above. In most cases where convergence proofs are provided, convexity and differentiability assumptions are needed. In few other cases, less restrictive assumptions suffice, such as the work of Zangwill [18] and Fiacco [7]. From a duality point of view, we can interpret the convergence results as resolution of duality gaps between the primal and dual problem.

In this paper we reinterpret the exterior penalty function method as a generalized lagrangian method which fills duality gaps in nonconvex programs. We present a simple duality theorem which can be implemented to devise a successful algorithm for solving the original problem (see [2]). Various geometric interpretations of the gap problem and its resolution are illustrated. In particular it is shown that penalty function methods may be viewed as linear supports in a modified "resource-objective" space.

2. GAP-FREE DUAL PROGRAMS

We will use the following notation: $\bar{g}(x) = \sum_{i=1}^{m} \bar{g}_i(x)$ and $\bar{h}(x) = \sum_{i=1}^{k} \bar{h}_i(x)$. Here $\bar{g}_i(x) = \max(0, g_i(x))$ and $\bar{h}_i$ may be defined either as $\bar{h}_i(x) = h_i^2(x)$ or as $\bar{h}_i(x) = |h_i(x)|$. Finally let $I(x) = \bar{g}(x) + \bar{h}(x)$.

Consider the following two primal and dual problems.

Problem P

Find $\inf \{f(x) : x \in S, g(x) \leqslant 0, h(x) = 0\}$.

Problem D

Find $\sup \theta(\lambda)$, where $\theta(\lambda) = \inf \{f(x) + \lambda I(x) = x \in S\}$.

Note that $I(x)$ measures the degree of infeasibility of x. For $x \in S$, $I(x)$ is zero if x is a feasible solution of Problem P and is positive otherwise. The following theorem establishes the fact that $\inf \{f(x) : x \in S, g(x) \leqslant 0, h(x) = 0\} = \sup_{\lambda \in R} \theta(\lambda) = \lim_{\lambda \to \infty} \theta(\lambda)$. Before proceeding any further it is worthwhile discussing the implications of the above equality. Since $\inf \{f(x) : x \in S, g(x) \leqslant 0, h(x) = 0\} = \lim_{\lambda \to \infty} \theta(\lambda)$ then we can get as close to the optimal value of problem P as we wish by picking a sufficiently large λ and solving the problem: minimize $f(x) + \lambda I(x)$. Of course, the optimal solution to the

above problem, x_λ, gets arbitrarily close to the feasible region, again by choosing λ large enough. It may be noted that other results similar to our theorem below are available in the literature of penalty functions (see for example [9, 17, 18]). A simple proof is given below for the sake of completeness.

THEOREM: Let S be compact and f, g, and h be continuous. Further suppose that problem P is consistent. Then $\inf \{(f(x) : x \in S, g(x) \leq 0, h(x) = 0\} = \sup_{\lambda \in R} \theta(\lambda) = \lim_{\lambda \to \infty} \theta(\lambda)$.

PROOF: The reader can easily verify the following statements:

i. $\inf \{f(x) : x \in S, g(x) \leq 0, h(x) = 0\} \geq \sup_{\lambda \in R} \theta(\lambda)$.

ii. For each λ there is an $x_\lambda \in S$ such that $\theta(\lambda) = f(x_\lambda) + \lambda I(x_\lambda)$.

iii. $\theta(\lambda)$ and $f(x_\lambda)$ are both nondecreasing functions of λ, whereas $I(x_\lambda)$ is a nonincreasing function of λ.

We will now show that $I(x_\lambda) \to 0$ as $\lambda \to \infty$. Let x be a feasible solution of problem P and let $\epsilon > 0$ be arbitrary. By choosing $\lambda = \frac{1}{\epsilon} | f(x) - f(x_1) | + 2$ we get $I(x_\lambda) \leq \epsilon$. Since $\lambda \geq 2 > 1$ if follows by iii above that $f(x_\lambda) \geq f(x_1)$. By contradiction suppose that $I(x_\lambda) > \epsilon$. By i and ii we get $\inf \{f(x) : x \in S, g(x) \leq 0, h(x) = 0\} \geq \theta(\lambda) = f(x_\lambda) + \lambda I(x_\lambda) > f(x_1) + \lambda\epsilon \geq f(x) + 2\epsilon > f(x)$. This is impossible, however, since x is a feasible solution of P. Therefore $I(x_\lambda) \leq \epsilon$ and by monotonicity of $I(x_\lambda)$ it is clear that $I(x_\lambda) \to 0$ as $\lambda \to \infty$. Consider the sequence $\{x_\lambda\}$ in S. By compactness of the latter there is a convergent subsequence, x_{λ_k} say, with limit $x_0 \in S$. By continuity of I (note continuity of g and h) we get $I(x_0) = I(\lim x_{\lambda_k}) = \lim I(x_{\lambda_k}) = 0$, i.e., x_0 is a feasible solution of Problem P. By continuity of f it is clear that $\lim f(x_{\lambda_k}) = f(x_0)$. Now $\sup_{\lambda \in R} \theta(\lambda) \geq \sup_k \theta(\lambda_k) = \lim_{k \to \infty} \theta(\lambda_k) = \lim_{k \to \infty} f(x_{\lambda_k}) + \lambda_k I(x_{\lambda_k}) \geq \lim_{k \to \infty} f(x_{\lambda_k}) = f(x_0) \geq \inf \{f(x) : x \in S, g(x) \leq 0, h(x) = 0\}$. In view of i the proof is complete.

From the above theorem it is clear that x_0 will solve the primal problem globally. The following points are worthwhile mentioning.

1. The importance of the compactness assumption of S is to be noted. This assumption will guarantee that for each λ, $\inf \{f(x) + \lambda I(x) : x \in S\}$ is achieved at some x_λ. This achievement is essential to our development and actually a duality gap may exist in the absence of this assumption as indicated by the numerical example given in the next section. It may be noted that this assumption is not very restrictive in most practical cases since the variables usually lie within certain bounds. Therefore we can pick the set S to be $\{x : | x_i | \leq \delta, i = 1, 2, \ldots, n\}$, where $\delta > 0$ is a sufficiently large number. Other technological constraints may be handled by $g(x) \leq 0$ and $h(x) = 0$. From a computational point of view, however, one may ignore the set S, i.e., minimize $f(x) + \lambda I(x)$ over R_n. If one gets into the trouble of nonexistence and nonconvergence of solutions (see discussion and example in section 3) then the set should be introduced.

2. The sequence x_λ is generally infeasible and if it becomes feasible at some point x'_λ, then x'_λ is a solution of the original problem. As shown in the proof of the theorem, the infeasibility of x_λ can be made arbitrarily small by choosing λ large enough.

3. The reader is warned against the false impression that the nonconvex Problem P can now be easily solved by minimizing $f(x) + \lambda I(x)$ over R_n (or S) with a large λ or with an increasing sequence of λ's. There is a number of difficulties associated with this approach. The first difficulty is that the resulting unconstrained problem is nonconvex and one is faced with the problem of local optimal solu-

tions. As a matter of fact the construction of $I = \bar{g} + \bar{h}$ may introduce new local minima. Another difficulty is that $\bar{g}_i$ and $\bar{h}_i$ are generally nondifferentiable even in the presence of differentiability of g_i and h_i. The last difficulty is the choice of λ. As the theorem indicates the larger the λ the closer we get to the optimal solution. However, one cannot afford to choose a single large λ since this places more emphasis on feasibility rather than optimality and usually causes slow convergence. On the other hand if one resorts to using an increasing sequence of λ's one faces the difficulty that a sequence of λ's may be appropriate for a certain problem but completely inappropriate for another problem. This makes a "rigid" rule of the form $\lambda_{k+1} = \alpha\lambda_k$, where α is a positive real number, not satisfactory. Even though various difficulties are associated with this approach, it is by no means computationally infeasible. After all, we are trying to solve a nonconvex problem to start with and no matter what general procedure is adopted one should expect some difficulties. In order to build a successful algorithm on the basis of the above theorem, some key measures are to be taken. First of all a nongradient type algorithm should be adopted. Secondly, an "interactive" type procedure is to be devised. By this is meant that one should be able to change the λ and also "perturb" or "bump" the current solution vector at different stages of the algorithm. This may be based on the progress in the objective and both the amount of infeasibility and the progress in the infeasibility measure. This does not only speed the rate of convergence but also decreases the likelihood of a local optimal solution. It is also advisable to use a different λ for each of the constraints. This will add the flexibility of being able to modify each penalty parameter according to the infeasibility in its constraint.

3. GEOMETRY OF DUALITY GAPS

In this section we will illustrate the notion of duality gaps and the notions of linear and nonlinear supports geometrically. Suppose we want to solve the problem: minimize $f(x)$ subject to $x \in S$ and $h(x) = 0$ (for simplicity's sake we delete the inequality constraint). Figure 1 represents the (h, f) plane where the point $(h(x), f(x))$ corresponds to $x \in S$. The collection of such points varying on $x \in S$ represents the shaded area.

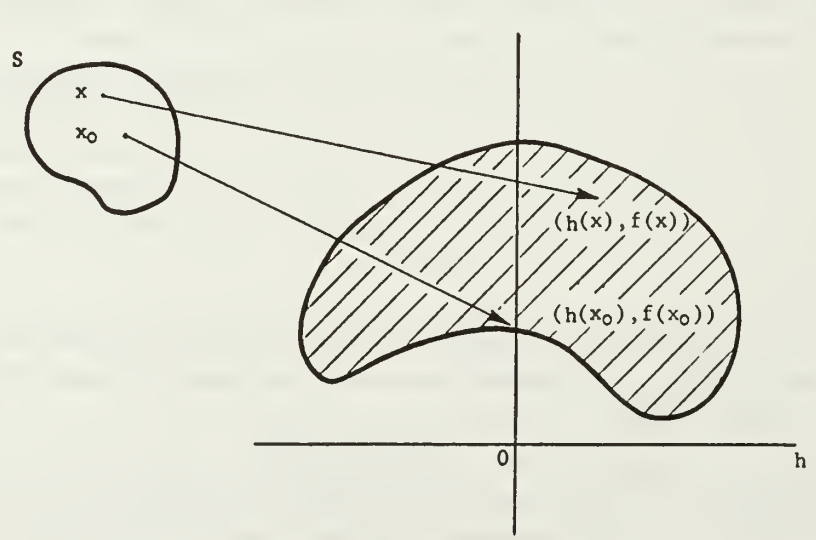


FIGURE 1. Duality gaps.

From Figure 1, it is clear that x_0 solves the problem since for any $x \epsilon S$ with $h(x) = 0$ we have $f(x) \geq f(x_0)$. As Everett [5] noted, one cannot find a μ such that the solution to the problem: minimize $f(x) + < \mu, h(x) >$ will also solve the equality problem. This is clear in Figure 2, since for any μ the minimization process of $f(x) + < \mu, h(x) >$ corresponds to a hyperplane that supports the shaded set. One can convince himself that we can never find a linear support of the set at the point $(h(x_0), f(x_0))$. As a matter of fact, one cannot find a linear support of the lower envelope of the set at a point having $h(x)$ between the numbers ϵ_1 and ϵ_2 of Figure 2. These points are referred to as being in a gap. For further discussion the reader may refer to [3, 5, 13, and 14].

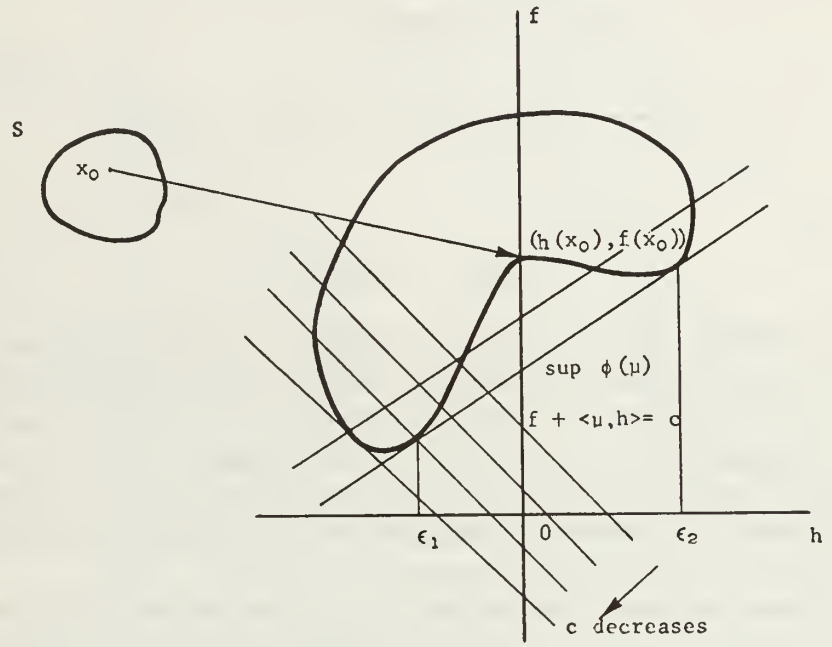


FIGURE 2. Linear supports.

As a matter of fact, we can illustrate the duality gap from a slightly different angle. The case shown in the diagram corresponds to a case when $\inf \{f(x) : x \epsilon S, h(x) = 0\} > \sup_{\mu} \phi(\mu)$, where $\phi(\mu) = \inf \{f(x) + < \mu, h(x) > : x \epsilon S\}$. Clearly $\phi(\mu)$ corresponds to the intercept of supporting lines with slope $-\mu$ on the f axis. It is obvious from Figure 2 that $\sup \phi(\mu)$ is strictly less than $f(x_0)$.

This also explains the fact that no duality gap exists in the case when f is convex, g is convex, h is linear, and S is convex. The reason is that the lower envelope of the shaded set will be convex. In other words, the function α corresponding to the lower envelope and defined by $\alpha(\epsilon_1, \epsilon_2) = \inf \{f(x) : x \epsilon S, g(x) \leq \epsilon_1, h(x) = \epsilon_2\}$ is convex. It is clear that if the lower envelope is convex then there is a supporting hyperplane at each boundary point, in particular at the point $(g(x_0), h(x_0), f(x_0))$. In other words, under suitable convexity assumptions, it is necessarily true that $\inf \{f(x) : x \epsilon S, g(x) \leq 0, h(x) = 0\} = \sup \{\phi(\lambda, \mu) : \lambda \geq 0\}$, where $\phi(\lambda, \mu) = \inf \{f(x) + < \lambda, g(x) > + < \mu, h(x) > : x \epsilon S\}$. We still have the problem of determining λ and μ which corresponds to the slope of the supporting hyperplane. See Figure 3 where the inequality constraint is deleted. Even under these strong convexity assumptions and the lack of a duality gap, one may never find (λ, μ) such that the solution of the problem: minimize $f(x) + < \lambda, g(x) > + < \mu, h(x) >$ will also solve our original problem. This is due to the fact that the supporting hyperplane may be vertical. This case corresponds to the case of lack

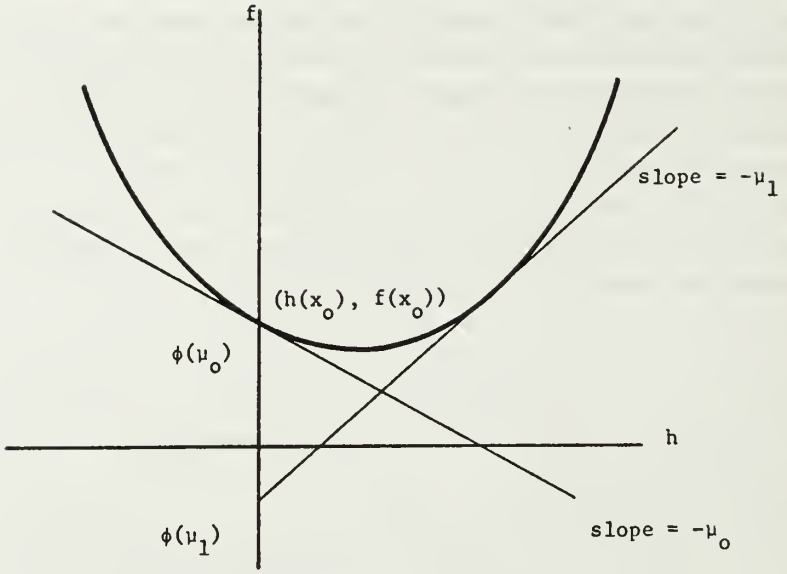


FIGURE 3. Convex envelope.

of a constraint qualification and only occurs if the lagrangian multiplier in the Fritz-John equation is zero. For further discussion on this point the reader may refer to Geoffrion [11] and Bazaraa [1].

So far, we have discussed the case of linear supports. As Figures 1 and 2 suggest, one can eliminate the gap by using nonlinear supports instead of linear supports. This idea is due to Gould [12] (see also Bellmore, Greenberg, and Jarvis [3]) who showed the existence of a nonlinear support. This is illustrated in Figure 4. The problem with this is that there is no way of telling the proper form of the supporting function for a given problem beforehand.

Now let us discuss the forms of support implied by the Theorem of section 2. We will refer to $\bar{g}$ as "truncated" support, $\bar{h} = |h|$ as an "absolute" support, and $\bar{h} = h^2$ as "parabolic" support. The reader

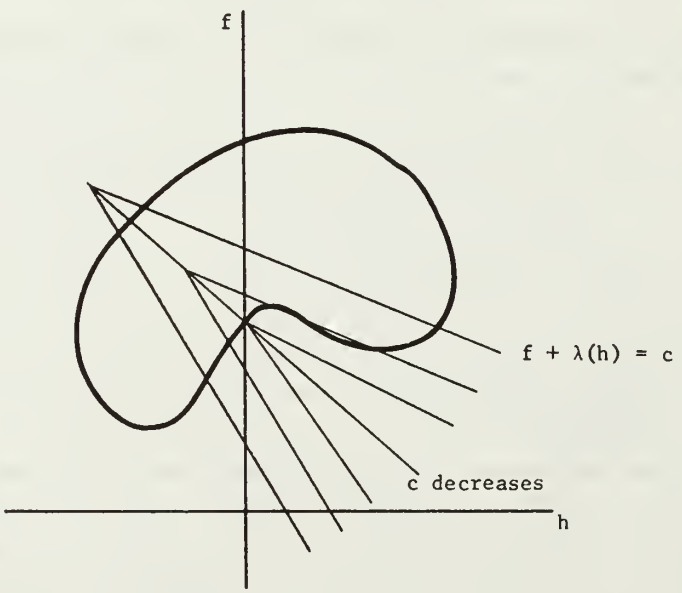


FIGURE 4. Nonlinear supports.

may note that these forms as well as others have been used by many authors, e.g., Fiacco and McCormick [9], Zangwill [17], Bellmore, Greenberg, and Jarvis [3], Greenberg [13], and Powell [15], to mention a few. Needless to say that other exterior penalty functions will close the duality gap if used in the duality theorem of the previous section. The "truncated," "absolute," and "parabolic" supports are used here for the clarity of presentation and for geometric appeal.

For the purpose of illustration we will discuss the cases of inequality and equality constraints separately. Figures 5 and 6 show the "truncated" support on the (g,f) plane and on the $(\bar{g},f)$ plane. It is clear that any "truncated" support with $\lambda \geq \lambda_1$ will close the duality gap, i.e., $\theta(\lambda)=f(x_0)$ for such

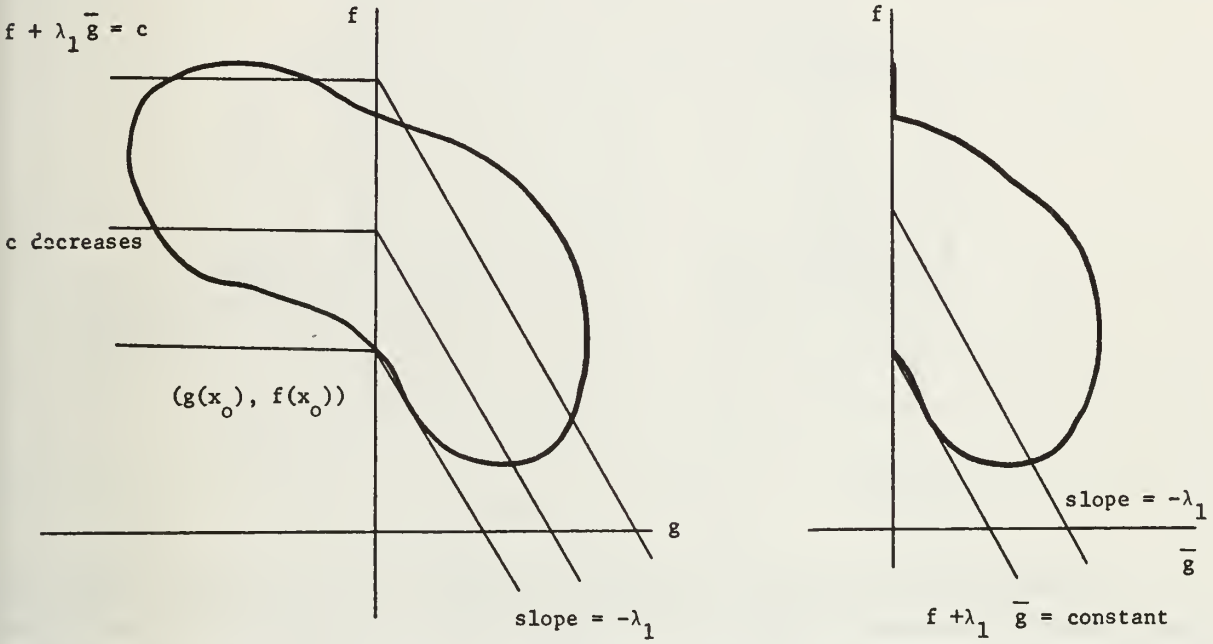


FIGURE 5. Truncated support, a boundary point.

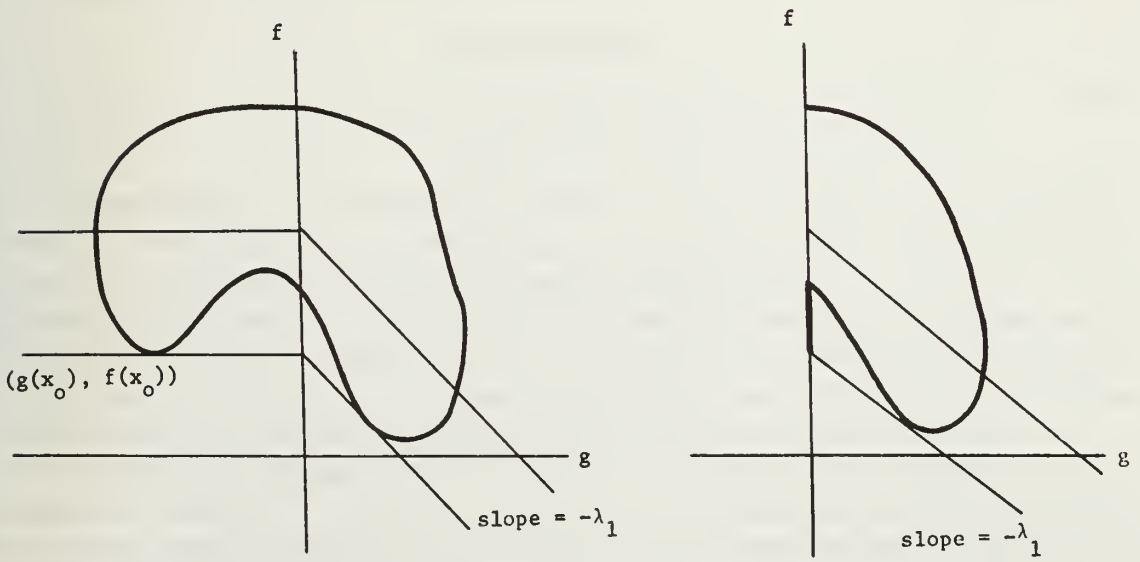


FIGURE 6. Truncated support, an interior point.

$\lambda \geq \lambda_1$. Note that the "truncated" support will appear as a linear support on the $(\bar{g}, f)$ plane where the region on the left of the f axis in the (g, f) plane is transformed onto the f axis on the $(\bar{g}, f)$ plane.

Figure 5 may be slightly altered such that the lower envelope of the $(\bar{g}, f)$ plane becomes tangential to the f axis at the point $(\bar{g}(x_0), f(x_0))$. In this case the duality gap will be closed only in a limiting sense, and for each λ we have $f(x_0) > \theta(\lambda)$. Figure 6 shows the case when $g(x_0) < 0$. Figure 7 depicts an "absolute" support (the case of a "parabolic" support is similar) for the case of equality constraint. The figure illustrates that one may be forced to use extremely large values of λ (i.e., steep supports) to dip close to the optimal solution.

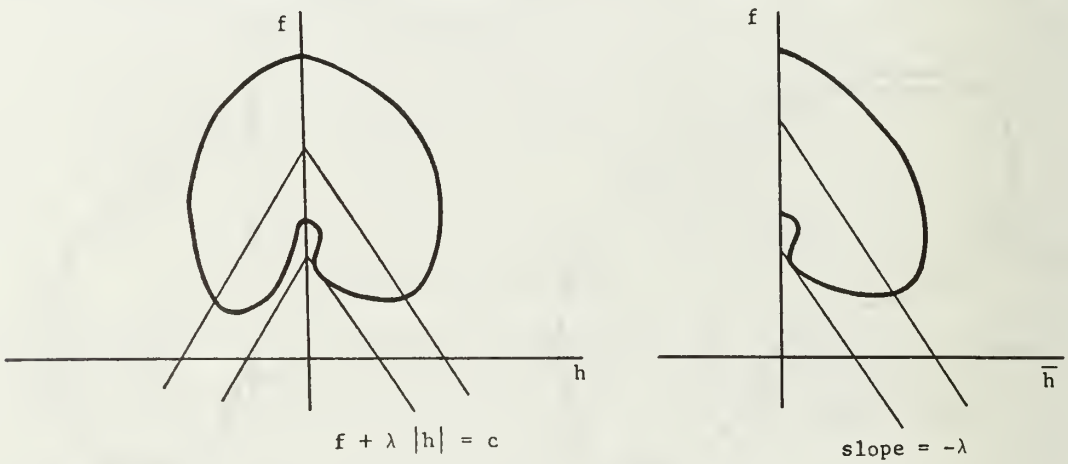


FIGURE 7. Absolute support.

As mentioned earlier one needs the compactness assumption of the set S in order to validate the duality theorem. The assumption insures that the infinum of $f(x) + \lambda I(x)$ over S is achieved at some x, which is essential in eliminating the duality gaps. The need for doing this is illustrated by the following example:

$$\text{minimize } x^3 + y^3$$

subject to

$$x + y - 1 = 0.$$

Note that the optimal solution to the above problem is ¼ and is achieved at the point (½, ½). Now let us find $\theta(\lambda) = \inf \{(x^3 + y^3) + \lambda(x + y - 1)^2\}$. It is clear that $\theta(\lambda) = -\infty$ for each fixed λ, e.g., by letting $x = y = -t$, where t is an arbitrarily large positive number. This shows that $\sup \theta(\lambda) = -\infty$ and a tremendous gap exists between the primal and dual problems. Before we attempt to solve this problem, it is worthwhile investigating this example on the (g, f) plane. Calculating the lower envelope function α defined by $\alpha(\epsilon) = \inf \{x^3 + y^3 : x + y - 1 = \epsilon\}$, we get $\alpha(\epsilon) = (1 + \epsilon)^3/4$ if $\epsilon \geq -1$ and $\alpha(\epsilon) = -\infty$ if $\epsilon < -1$. This is illustrated in Figure 8. Note that for each λ, we can never find a support of the shaded set. No matter how steep our parabola (or "absolute" support) is, somewhere along the line we will hit the vertical line passing through $\epsilon = -1$. This explains why $\theta(\lambda) = -\infty$ for each λ and is precisely the reason why we need compactness of the set S. Now if we introduce the bounds on the variables we will be able to cut the lower portion of the cylinder to the left of $\epsilon = -1$ and accordingly we can

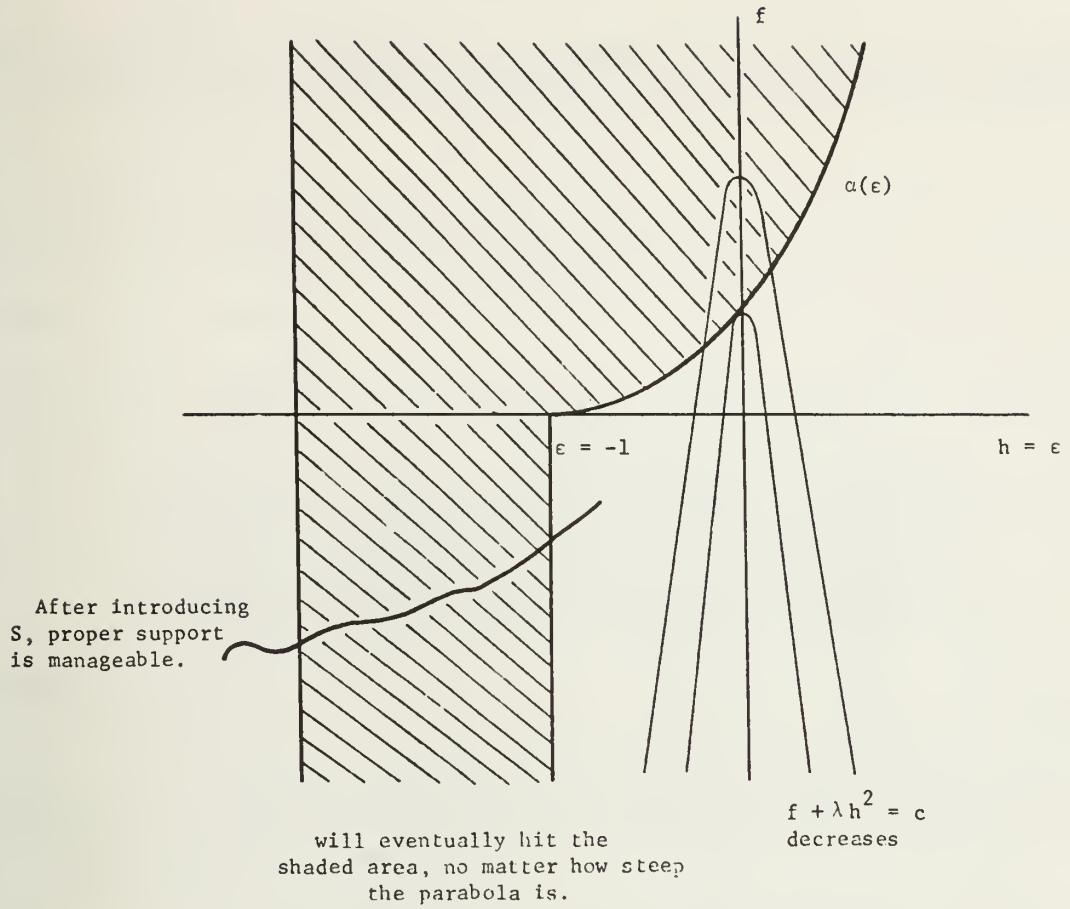


FIGURE 8. Compactness assumption.

support the shaded set at points close to the optimal solution. Therefore, the set S will indeed close the above gap between the dual problems.

REFERENCES

[1] Bazaraa, M. S., "A Theorem of the Alternative with Application to Convex Programming: Optimality, Duality, and Stability," to appear in J. Math. Analysis and Applications.

[2] Bazaraa, M. S. and R. Rardin, "A Modified Rosenbrock Procedure for Solving Constrained Problems," Georgia Institute of Technology, (Sept. 1972).

[3] Bellmore, M., H. J. Greenberg, and J. J. Jarvis, "Generalized Penalty Function Concepts in Mathematical Optimization," Operations Research, vol. 18 (1970).

[4] Brooks, R., and A. Geoffrion, "Finding Everett's Lagrange Multipliers by Linear Programming," Operations Research, vol. 14 (1966).

[5] Everett, H., "Generalized Lagrange Multiplier Method for Solving Problems of Optimum Allocation of Resources," Operations Research, vol. 11 (1963).

[6] Falk, J. E., "Lagrange Multipliers and Nonlinear Programming," J. Math. Analysis and Applications, vol. 19 (1967).

[7] Fiacco, A. V., "Penalty Methods for Mathematical Programming in E^n with General Constraint Sets," J. Optimization Theory and Applications, vol. 6 (1970).

[8] Fiacco, A. V. and G. P. McCormick, "Extensions of SUMT for Nonlinear Programming: Equality Constraints and Extrapolation," Management Science, vol. 12 (1966).

[9] Fiacco, A. V. and G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques* (John Wiley and Sons, New York, 1968).

[10] Fiacco, A. V. and G. P. McCormick, "The Sequential Unconstrained Minimization Technique for Nonlinear Programming, A Primal-Dual Method," Management Science, vol. 10 (1964).

[11] [illegible], *Nonlinear Programming: A Simplified Applications-Oriented Development* [illegible] (1967).

[12] Gould, F. J., "Extensions of Lagrange Multipliers in Nonlinear Programming," SIAM J. Applied Math., vol. 17 (1969).

[13] Greenberg, H. J., "Lagrangian Duality Gaps: Their Source and Resolution," Technical Report CP-69005, Southern Methodist University (1969).

[14] [illegible], "[illegible] — A Heuristic Presentation," OEG Research Number 76, [illegible], Washington, D.C. (1956).

[15] Powell, M. J. D., "A Method for Nonlinear Constraints in Minimization Problems," Conference on Optimization, Institute of Mathematics and Its Applications, 283–297 (1969).

[16] Rockafellar, R. T., *Convex Analysis* (Princeton University Press, Princeton, 1970).

[17] Zangwill, W. I., *Nonlinear Programming: A Unified Approach* (Prentice-Hall, New Jersey, 1969).

[18] Zangwill, W. I., "Nonlinear Programming via Penalty Functions," Management Science, vol. 13 ([illegible]).

CONVERGENCE OF THE BOUNDED FIXED CHARGE PROGRAMMING PROBLEM*

Charles I. Tiplitz

Logistics Management Institute
Washington, D.C.

ABSTRACT

Calculating the solution for a mixed integer linear programming problem has been problematic for any sizable dimension because of the time required. Accordingly, an improved method for the fixed charge problem is presented here. The method is a modification of a bounding technique first suggested by Balinski, and it exploits fully the ratio of cost-to-use, first described by Cooper and Drebes. It exploits those local properties which have global application so that enumeration is confined to those relatively few combinations which cannot be evaluated otherwise.

I. INTRODUCTION

The fixed charge problem is a special form of the mixed integer linear programming problem. While the mixed integer linear programming (MILP) problem, if feasible, has a solution, calculating it has been problematic for any sizable dimension because of the time required (see [5]). Accordingly, a method for the fixed charge problem is suggested here.

In the ordinary linear programming problem, the property of convex cones provides clues about global properties from local explorations. Hence the search for solutions can be narrowed, eliminating the need for an enumerative search. In integer problems, this advantage is not readily available. The method suggested here is a modification of a bounding technique first suggested by Balinski [1] and it exploits fully a ratio described by Cooper and Drebes [2]. It will result, as will be seen, in exploitation of those local properties which have global application so that enumeration is confined to those relatively few combinations which cannot be evaluated otherwise. Thus a global optimum may be determined in far fewer iterations than by branch and bound and in fewer iterations than the ranking method of Murty [4] which reduces, but does not minimize the extreme points to be considered.

II. PROBLEM DEFINITION

The fixed charge problem commonly arises in an industrial situation where alternative methods for manufacturing several or more products exist. The cost to manufacture each such product includes a charge proportional to the number of items made and a fixed charge, usually for tooling or to open up a manufacturing line, regardless of the quantity produced. The fixed charge problem may be formulated as follows:

$$\text{minimize } \phi = \sum f_j(x_j),$$

*This work was done in connection with a study by the Logistics Management Institute for the Assistant Secretary of Defense (Installations and Logistics).

subject to

$$Ax = b$$

$$x \geqslant 0,$$

where

$$f_j(x_j) = c_j x_j + d_j \delta_j,$$

$$\delta_j = \begin{cases} 0 & \text{if } x_j = 0 \\ 1 & \text{if } x_j > 0. \end{cases}$$

In many fixed charge problems and certainly most industrial ones, there is a capacity or upper bound on each variable which is imbedded in $Ax = b$. For example, the capacity of an assembly line may be a_{11} units of product #1 or a_{12} units of product #2. This might be represented by the equation:

$$x_1/a_{11} + x_2/a_{12} \leqslant 1,$$

where x_1 represents the number of units of product #1 assembled. Somewhere else in matrix A will be found the requirement for product #1, say:

$$x_1 \geqslant b_1.$$

Finally, operation of the assembly line at a nonzero level may be the source of one of the fixed charges. For the purpose at hand it is convenient to rearrange these facts by redefining x_i to represent the proportion of capacity devoted to assembly of product i. Then the two equations would be:

$$x_1 + x_2 \leqslant 1$$

and

$$a_{11}x_1 \geqslant b_1.$$

If we create a new variable, $y = 0$ or 1, to represent the operation of the assembly line, then we can further change the first equation:

$$x_1 + x_2 - y_1 \leqslant 0.$$

Also we may now assign the corresponding fixed costs as the coefficient of the new variable in a modified objective function.

Thus, for many problems, we may consider the bounded fixed charge problem as follows:

PROBLEM I:

$$\text{Minimize } \Phi'(x, y) = \sum_i c_i x_i + \sum_j d_j y_j,$$

subject to

$$y_j = 0 \text{ or } 1,$$

$$x_i \geq 0,$$

$$c_i,\ d_j \geq 0,$$

$$Ax + By = e,$$

where the last equation is such that for each j and for each subset of indices of x, s_j, $\sum_{i \in s_j} x_i \leq y_j$, where $s_j \cap s_k = \phi$ if $j \neq k$.

We have constrained the x_j to be one or less for computational convenience, although no generality is lost thereby. An example of the matrix structure is given in Table I.

TABLE I. *Illustrative matrix*

$x_1 + x_2$				$-y_1$				≤ 0
	x_3				$-y_2$			≤ 0
		x_4	$+x_5$			$-y_3$		≤ 0
$a_{41}x_1$	$+a_{43}x_3$							$\geq e_4$
	$a_{52}x_2$	$+a_{54}x_4$	$+a_{55}x_5$					$\geq e_5$

III. BOUNDING

The bounding technique of Balinski is carried over into this problem by allowing y_j to be noninteger as follows:

PROBLEM II:

$$\text{Minimize } \phi^{II}(x, y) = \sum_i c_i x_i + \sum_j d_j y_j$$

subject to the same constraints as Problem I except:

$$0 \leq y \leq 1.$$

Let x^I, y^I be the solution to Problem I, and x^{II}, y^{II} be the solution to Problem II. Since x^I, y^I are feasible in Problem II and x^{II}, y^{II} may not be feasible in Problem I, it follows that:

$$\phi^I(x^I, y^I) = \phi^{II}(x^I, y^I) \geq \phi^{II}(x^{II}, y^{II}).$$

Therefore, the solution of Problem II is a lower bound to the solution of Problem I. Note also that:

$$\phi^u = \sum_i c_i x_i^{II} + \sum_j d_j \delta_j^{II} \geq \phi^I(x^I, y^I),$$

where

$$\delta_j^{II} = 1 \qquad \text{if } y_j^{II} \neq 0$$

and

$$\delta_j^{II} = 0 \qquad \text{if } y_j^{II} = 0.$$

ϕ^u is an upper bound in the sense that if there is a better solution, say ϕ^a, then $\phi^a < \phi^u$; however, a lower upper bound may sometimes be obtained by solving the noninteger linear program problem:

PROBLEM III:

$$\text{Minimize } \phi^{III}(x) = \sum_i c_i x_i^{III} + \sum_j d_j \delta_j^{II},$$

where

$$Ax = e - B\delta^{II}.$$

It follows that $\phi^{III}(x^{III}) \leq \phi^u$, as desired.

IV. CONVERGENCE PROCEDURE

So far only bounds have been obtained. We now introduce the local criterion which enable us to infer certain global implications. The criterion is essentially a method for weighting the fixed charges. Therefore, consider next Problem IV, derived from Problem II by adjusting d_j as follows:

If $y_j \neq 0$ in y^{II}, let $d_j^1 = \dfrac{d_j}{y_j^{II}}$, then:

PROBLEM IV:

$$\text{Minimize } \Phi^{IV}(x, y, d_j^1) = \sum_i c_i x_i + \sum_j d_j^1 y_j,$$

subject to the constraints of Problem II.

By construction: $\Phi^u = \Phi^{IV}(x^{II}, y^{II}, d_j^1) \geq \Phi^{IV}(x^{IV}, y^{IV}, d_j^1) \geq \Phi^{II}(x^{II}, y^{II})$. The procedure may be iterated. That is, we may let:

$$d_j^2 = \frac{d_j}{y_j^{IV}} \text{ and minimize } \Phi^{IV} \text{ again.}$$

Then one of two alternatives occurs: At some iteration, say at k, $d_j^k = d_j^{k-1}$ for all j, i.e., we have convergence; or at every iteration there is some j for which $d_j^k \neq d_j^{k-1}$. The latter will occur if for some j and k, $y_j^k = 0$ and $y_j^{k+1} \neq 0$ or $y_j^k \neq 0$ and $y_j^{k+1} = 0$. In either case, $d_j^k \neq d_j^{k-1}$, that is, $y_j^k \neq y_j^{k-1}$. Otherwise, d_j is monotonically increasing and convergence will occur.

Convergence may also be too slow and require more iterations than desired to reach prescribed precision. Such variables may be treated as if nonconvergent. Nonconvergence is discussed in section VI. In either event, successive iterations should, for efficiency, use the previous solution as a starting basis.

V. CONVERGENCE CONDITIONS

We will first examine convergent solutions in the simple cases and then discuss convergent solutions in the general situation. In section VI we will discuss treatment of nonconvergent variables.

TABLE II. *First illustrative example*

$$\begin{aligned} x_1 \qquad\qquad - y_1 \qquad\qquad &\leqslant 0 \\ x_2 \qquad\qquad - y_2 \qquad\qquad &\leqslant 0 \\ a_{31}x_1 + a_{32}x_2 \qquad\qquad &\geqslant e_3 \\ \text{Minimize } \Phi = c_1x_1 + c_2x_2 + d_1y_1 + d_2y_2 & \end{aligned}$$

Compare first the cost coefficients of Problem IV for the example illustrated in Table II when there is convergence and $e_3 < a_{31}$, $e_3 < a_{32}$. Since the optimum is at an extremum either $x_1^{IV} = 0$ or $x_2^{IV} = 0$. If $x_1^{IV} = 0$, then from solution of Problem IV,

$$\frac{c_1}{a_{31}} + \frac{d_1}{e_3} > \frac{c_2}{a_{32}} + \frac{d_2}{e_3}\,.$$

Otherwise the inequality is reversed. Now if the two alternate values $x_1^I = 0$ and $x_2^I = 0$ are alternatively substituted in the matrix A of Table II, it will be seen that the values of the objective function of Problem I have the same inequalities as that of Problem IV just given. In other words, under the above circumstances, the solution to Problem IV is the solution to Problem I.

Consider next the example in Table III. If $x_1^{IV} = 0$, and $a_{33} > e_3$, then by direct computation:

$$\frac{c_1e_3}{a_{31}} + \frac{c_2e_4}{a_{42}} + d_1 > \frac{c_2e_4}{a_{42}} + \frac{c_3e_3}{a_{33}} + d_1 + d_2,$$

which is identical to Problem I. Again we may substitute Problem IV for Problem I.

TABLE III. *Second illustrative example*

$$\begin{aligned} x_1 + x_2 \qquad - y_1 \qquad\qquad &\leqslant 0 \\ x_3 \qquad\qquad - y_2 \qquad &\leqslant 0 \\ a_{31}x_1 + a_{33}x_3 \qquad\qquad &\geqslant e_3 \\ a_{42}x_2 \qquad\qquad &\geqslant e_4 \\ \text{Minimize } \Phi = c_1x_1 + c_2x_2 + c_3x_3 + d_1y_1 + d_2y_2 & \end{aligned}$$

Even in simple problems it is easy to find cases when the solution to Problem IV will not furnish the solution to Problem I. Consider, for example, that in Table II, $e_3 > a_{32}$. Total costs (Problem I) if $x_2 = 0$ are:

$$\Phi_a = \frac{c_1e_3}{a_{31}} + d_1;$$

and if $x_2 = 1$

$$\Phi_b = c_2 + d_2 + c_1\left(\frac{e_3 - a_{32}}{a_{31}}\right) + d_1.$$

Then

$$\Delta\Phi = \Phi_b - \Phi_a = c_2 + d_2 - \frac{c_1a_{32}}{a_{31}}\,.$$

If $\Delta\Phi > 0$, i.e., $x_2^I = 0$, then

$$\frac{c_1}{a_{31}} < \frac{c_2}{a_{32}} + \frac{d_2}{a_{32}}.$$

If, however, $x_2^{IV} = 1$, then from the simplex procedure:*

$$\frac{c_2}{a_{32}} + \frac{d_2}{a_{32}} < \frac{c_1}{a_{31}} + \frac{d_1}{e_3}.$$

Combining the last two inequalities we obtain:

$$\frac{c_1}{a_{31}} < \frac{c_2}{a_{32}} + \frac{d_2}{a_{32}} < \frac{c_1}{a_{31}} + \frac{d_1}{e_3}.$$

Thus in this simple case, to detect the existence of necessary conditions for convergence of the solution of Problem IV to the wrong solutions of Problem I, one need only verify the left-hand inequality which can be done by recalculating the relative cost coefficients with $d_j = 0$ when $d_j \neq d_j^k$. This is equivalent to the question: what vectors would be forced out of the solution if $d_j = 0$.

In the general case, the matrix $A + B$ is identical for both Problem I and Problem IV. Except as in the above cases, the equations for Φ will lead to identical inequalities and $x^I = x^{IV}$. Necessary conditions for convergence to the wrong solution, as in the simple cases, can be readily detected in the general case. The proof follows the procedure above and is given in the appendix. The criterion is the same: For all j such that $0 < y_j < 1$, let $d_j = 0$ and list the vectors to be forced out. This can be determined simply by picking out the dual cost coefficients of the slack nonbasic, noninteger variables in Problem IV which become negative when $d_j = 0$. If there are none, then a global optimum has been obtained and, moreover, the conditions for such an optimum are readily identified. Otherwise, enumeration or branch and bound are necessary for the usually relatively few variables involved.

VI. ENUMERATION

If all variables in Problem IV do not converge (or some variable converges too slowly), the problem appears more formidable. Hence, from Murphy's law we may expect the nonconvergent case to be more common. Fortunately, most often the $y_i = 0$, or 1 even when allowed to take noninteger values and most of the remainder will not oscillate. Hence, we may be able to examine enumeratively, with relative ease, the relatively few combinations of only the oscillating or nonconvergent variables. The method of branch and bound can be used, for example. Initial bounds are also provided from Problems II–IV, to speed up the process.

VII. RESULTS

In one small fixed charge problem (29 integers and 71 continuous variables and the matrix was 8 percent filled), convergence to a global optimum occurred. About one-fifth of the integer variables were noninteger in the Problem IV format. In a larger problem (172 integers and 553 continuous variables and the matrix was 2 percent filled), the number of noninteger y_j in the Problem IV format was 27. Of these it appeared that two would oscillate and eight converged incorrectly. Thus full exploration was completed with ease so that a global optimum was simple to identify.

*This procedure may be found in a variety of sources. For example, see [3].

The ratios of fixed to variable costs (about 1 : 11 for the large problem and 1 : 5 for the small problem) are probably typical of many other multi-year industrial planning situations. Under those circumstances I would expect that relatively few variables would oscillate or not converge. As the ratio increases, the problem begins to resemble a pure integer problem and the technique loses its advantage as a smaller and smaller proportion of variables converge.

Appendix

First we will consider two adjacent feasible solutions, one of which is the solution to Problem I and the other solves Problem IV. We will derive the condition for this misleading convergence in Problem IV and then discuss this condition for more general circumstances, e.g., nonadjacent solutions, by-products.

This value for the objective function for the two adjacent feasible solutions is

$$Z_1^I = \sum_j \left(\sum_{i \epsilon s_j} c_i x_i^1 + d_j \right) + c_m x_m^1 + d_m \tag{1}$$

and

$$Z_2^I = \sum_j \left(\sum_{i \epsilon s_j} c_i x_i^2 + d_j \right) + c_n x_n^2 + d_n + d_m. \tag{2}$$

Assume that

$$Z_1^I < Z_2^I, \tag{3}$$

but that Problem IV has converged to x^2. That is, the simplex cost criterion, viz. the relative cost coefficient of x_m, is positive:

$$c_m + \sum_p a_{pm} \pi_p > 0, \tag{4}$$

where π_p are the simplex multipliers given by [3]:

$$\pi_p = -\sum_i \beta_{ip} c_i, \tag{5}$$

where β_{ip} are the elements of the inverse of the matrix in the basic solution of Problem IV and the c_i are the cost coefficients of the basic vectors in this solution, some of which are the d_j^k.

The relation between the vectors x_i and x_n in Equations (1) and (2) can be given as follows [3]:

$$x_i^1 = x_i^2 - x_n^2 \frac{\sum_p \beta_{ip} a_{pm}}{\sum_p \beta_{np} a_{pm}}, \qquad i \neq m. \tag{6}$$

Putting Equations (1), (2), and (6) in (3), we obtain

$$c_m x^1_m < d_n + c_n x^2_n + x^2_n \frac{\sum_{i \neq m,\, n} c_i \sum_p \beta_{ip} a_{pm}}{\sum_p \beta_{np} a_{pm}}. \tag{3a}$$

Initially we assume that in the matrix $A + B$ there is only one row in which the coefficients of the vectors x_m and x_n are both nonzero. This assumption is equivalent to assuming no byproducts and no component requirements. Then, using the fact from Equation (6), that $x^1_i = x^2_i$ except when $i = n$, or m, we deduce that:

$$\sum_p \beta_{ip} a_{pm} = 0, \qquad i \neq m,\ n. \tag{7}$$

Therefore, the last term in Equation (3a) vanishes.

Returning now to the simplex criterion of Equation (4), substitute Equations (5) and (7) and relabel the c_i where appropriate, and we obtain

$$c_n \sum_p \beta_{np} a_{pm} + d^k_n \sum_p \beta_{sp} a_{pm} < c_m - d^k_m \sum_p \beta_{tp} a_{pm}, \tag{4a}$$

where s and t refer to the columns in which y_n and y_m appear. The relation between x_m and x_n in Equations (1) and (2) is [3]:

$$x^1_m = \frac{x^2_n}{\sum_p \beta_{np} a_{pm}}. \tag{8}$$

Using this, the formula for d^k_n in Problem IV, and direct computation yields:

$$\frac{d_n}{x^1_m} + \frac{c_n x^2_n}{x^1_m} < c_m + \frac{d_m}{x^1_m}. \tag{4b}$$

Equations (3a) and (4b) now result in:

$$c_m x^1_m < c_n x^2_n + d_n < c_m x^1_m + d_m, \tag{9}$$

which is the desired form.

If we relax our initial assumption and allow two or more rows in which coefficients of x_m and x_n are both nonzero, Equation (9) still applies since all parts of the inequalities are affected equally by the change in Equation (7). Also, if we consider nonadjacent solutions, Equation (9) is applicable to each pair of vectors. Hence the conditions for misleading convergence may be tested as before.

REF[illegible]

[1] Balinski, M. L., "Computing [illegible]", Navy Supply System Res[illegible]

[2] Cooper, L. and C. Drebes, [illegible]," Nav. Res. Log. Quart. *14*, 1[illegible]

[3] Garvin, W. W., Introduction [illegible]

[4] Murty, K. G., "Solving the Fixe[illegible]" Operations Research *16*, 268–279 (Ma[illegible])

[5] Trauth, C. A., Jr., and R. E. W[illegible] Efficiency," Management S[illegible]

U.S. GOVERNMENT PRINTING OFFICE [illegible]

INFORMATION FOR CONTRIBUTORS

The NAVAL RESEARCH LOGISTICS QUARTERLY is devoted to the dissemination of scientific information in logistics and will publish research and expository papers, including those in certain areas of mathematics, statistics, and economics, relevant to the over-all effort to improve the efficiency and effectiveness of logistics operations.

Manuscripts and other items for publication should be sent to The Managing Editor, NAVAL RESEARCH LOGISTICS QUARTERLY, Office of Naval Research, Arlington, Va. 22217. Each manuscript which is considered to be suitable material for the QUARTERLY is sent to one or more referees.

Manuscripts submitted for publication should be typewritten, double-spaced, and the author should retain a copy. Refereeing may be expedited if an extra copy of the manuscript is submitted with the original.

A short abstract (not over 400 words) should accompany each manuscript. This will appear at the head of the published paper in the QUARTERLY.

There is no authorization for compensation to authors for papers which have been accepted for publication. Authors will receive 250 reprints of their published papers.

Readers are invited to submit to the Managing Editor items of general interest in the field of logistics, for possible publication in the NEWS AND MEMORANDA or NOTES sections of the QUARTERLY.

NAVAL RESEARCH
LOGISTICS
QUARTERLY

JUNE 1973
VOL. 20, NO. 2
NAVSO P-1278

CONTENTS

ARTICLES

OFFICE OF NAVAL RESEARCH
Arlington, Va. 22217